GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 26, 2003, 10:11:26 ; Search time 1905 Seconds
(without alignments)
10136.138 Million cell updates/sec

Title: US-10-045-116-1
Perfect score: 472
Sequence: 1 AGCCACCACCCAGTGAGCCT.....ACGTTCCTGGAACTCAAGCA 472

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		471	99.8	3500	6	AX705342	AX705342 Sequence
2		471	99.8	11288	6	AR134885	AR134885 Sequence
3		471	99.8	11288	6	AR182304	AR182304 Sequence
4		471	99.8	11288	6	AX269130	AX269130 Sequence
5		471	99.8	11288	9	HSCRCANTA	Z21818 H.sapiens c
6		471	99.8	15056	6	AR266427	AR266427 Sequence
7		471	99.8	15056	6	AX259954	AX259954 Sequence
8		471	99.8	15056	6	AX262359	AX262359 Sequence
9		471	99.8	39707	9	AC008999	AC008999 Homo sapi
10		455.8	96.6	4137	9	HSCEATG	X62151 H.sapiens c
11		432.4	91.6	3281	6	A37261	A37261 Sequence 1
12		432.4	91.6	3281	9	HUMCEA01	M59255 Human carci
13		278.2	58.9	3500	6	AX705383	AX705383 Sequence
14		271.8	57.6	3500	6	AX705405	AX705405 Sequence
15		264.8	56.1	645	9	HSCGM2PRO	X98312 H.sapiens c
16	c	263.2	55.8	39753	9	AC005797	AC005797 Homo sapi
17	c	263	55.7	3500	6	AX705384	AX705384 Sequence
18	c	256.6	54.4	3500	6	AX705406	AX705406 Sequence
19		234	49.6	414	9	HUMCEAAB2	M60963 Human carci
20		218.6	46.3	39233	9	AC005626	AC005626 Homo sapi
21		158.2	33.5	40520	9	AC011513	AC011513 Homo sapi
22		156.6	33.2	2073	9	HUMCEANCA	M17082 Human carci
23	c	134	28.4	38235	9	AC004559	AC004559 Homo sapi
24		128.8	27.3	2450	9	HSACGM1	AF110325 Homo sapi
25	c	120.8	25.6	39801	9	AC005955	AC005955 Homo sapi
26	c	120.8	25.6	44332	9	AC005794	AC005794 Homo sapi
27	c	104.4	22.1	243197	9	AC005392	AC005392 Homo sapi
28		104.2	22.1	3706	9	HSBGPEX1	X67277 H.sapiens B
29	c	103.2	21.9	39823	9	AC004785	AC004785 Homo sapi
30	c	97.6	20.7	173298	2	AC118918	AC118918 Rattus no
31	c	97.6	20.7	262466	2	AC095945	AC095945 Rattus no
32		94	19.9	3774	6	AR134886	AR134886 Sequence
33		94	19.9	3774	6	AR182305	AR182305 Sequence
34		94	19.9	3774	9	HSU17131	U17131 Human carci
35		94	19.9	39753	9	AC005797	AC005797 Homo sapi
36		94	19.9	218996	2	AC140257	AC140257 Mus muscu
37		94	19.9	226861	2	AC134475	AC134475 Mus muscu
38		86.6	18.3	43593	9	AC022516	AC022516 Homo sapi
39		85.8	18.2	39801	9	AC005955	AC005955 Homo sapi
40		82.2	17.4	2690	9	HUMCEAB	M22433 Homo sapien
41		81.8	17.3	221111	2	AC079526	AC079526 Mus muscu
42		80	16.9	1482	10	MUSPGGLYB1	M83340 Mouse pregn
43	c	80	16.9	223585	2	AC079497	AC079497 Mus muscu
44		78.4	16.6	237000	2	AC138311	AC138311 Mus muscu
45		77.4	16.4	288952	2	AC106470	AC106470 Rattus no

ALIGNMENTS

RESULT 1
AX705342
LOCUS AX705342 3500 bp DNA linear PAT 04-APR-2003
DEFINITION Sequence 11 from Patent WO03014388.
ACCESSION AX705342
VERSION AX705342.1 GI:29562007
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Distler,J., Model,F. and Taubert,H.
TITLE Method and nucleic acids for the analysis of colon cancer
JOURNAL Patent: WO 03014388-A 11 20-FEB-2003;

```
           Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .3500
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT       843 a   1039 c    931 g    687 t
ORIGIN

  Query Match             99.8%;  Score 471;  DB 6;  Length 3500;
  Best Local Similarity   100.0%;  Pred. No. 6.3e-141;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2506 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 2565

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2566 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 2625

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2626 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 2685

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2686 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 2745

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2746 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 2805

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2806 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 2865

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2866 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 2925

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2926 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 2976


RESULT 2
AR134885
LOCUS       AR134885               11288 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6194211.
ACCESSION   AR134885
VERSION     AR134885.1  GI:14123790
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11288)
  AUTHORS   Richards,C.Ann. and Huber,B.
  TITLE     Transcriptional regulatory sequence of carcinoembryonic antigen for
            expression targeting
  JOURNAL   Patent: US 6194211-A 1 27-FEB-2001;
FEATURES             Location/Qualifiers
     source          1. .11288
                     /organism="unknown"
BASE COUNT       2944 a   3060 c   2955 g   2329 t
ORIGIN

  Query Match             99.8%;  Score 471;  DB 6;  Length 11288;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-141;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2926 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 2985

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2986 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 3045

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3046 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 3105

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3106 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 3165

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3166 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 3225

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3226 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 3285

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3286 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 3345

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3346 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 3396


RESULT 3
AR182304
LOCUS       AR182304               11288 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 4 from patent US 6337209.
ACCESSION   AR182304
VERSION     AR182304.1  GI:20225220
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 11288)
  AUTHORS   Huber,B. and Richards,C.A.
  TITLE     Molecular constructs containing a carcinoembryonic antigen
            regulatory sequence
  JOURNAL   Patent: US 6337209-A 4 08-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .11288
                     /organism="unknown"
BASE COUNT       2944 a   3060 c   2955 g   2329 t
ORIGIN

  Query Match             99.8%;  Score 471;  DB 6;  Length 11288;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-141;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3346 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 3405

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3406 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 3465

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3466 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 3525

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3526 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 3585

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10534 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 10593

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10594 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 10653

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10654 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 10713

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10714 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 10764


RESULT 4
AX269130
LOCUS       AX269130               11288 bp    DNA     linear   PAT 29-OCT-2001
DEFINITION  Sequence 1 from Patent WO0174861.
ACCESSION   AX269130
VERSION     AX269130.1  GI:16542049
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Vile,R.G., Harrington,K., Murphy,S. and Bateman,A.
  TITLE     Compositions and methods for tissue specific gene regulation
            therapy
  JOURNAL   Patent: WO 0174861-A 1 11-OCT-2001;
            MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .11288
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             10804. .10938
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD10314.1"
                     /db_xref="GI:16542050"
                     /translation="MESPSAPPHRWCIPWQRLLLTGEGRTTWERVGGGSWGLLGRTGL
                     "
BASE COUNT     2944 a   3060 c   2955 g   2329 t
ORIGIN

  Query Match           99.8%;  Score 471;  DB 6;  Length 11288;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-141;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10294 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 10353

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10354 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 10413

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10414 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 10473

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10474 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 10533

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10534 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 10593

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10594 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 10653

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10654 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 10713

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10714 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 10764


RESULT 5
HSCRCANTA
LOCUS       HSCRCANTA              11288 bp    DNA     linear   PRI 29-MAY-1996
DEFINITION  H.sapiens carcinoembryonic antigen gene.
ACCESSION   Z21818
VERSION     Z21818.1  GI:437783
KEYWORDS    carcinoembryonic antigen.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schrewe,H., Thompson,J., Bona,M., Hefta,L.J., Maruya,A.,
            Hassauer,M., Shively,J.E., von Kleist,S. and Zimmermann,W.
  TITLE     Cloning of the complete gene for carcinoembryonic antigen: analysis
            of its promoter indicates a region conveying cell type-specific
            expression
  JOURNAL   Mol. Cell. Biol. 10 (6), 2738-2748 (1990)
  MEDLINE   90258861
   PUBMED   2342461
  REMARK    (sites)
REFERENCE   2  (bases 1 to 11288)
  AUTHORS   Richards,C.A., Wolberg,A.S. and Huber,B.E.
  TITLE     The transcriptional control region of the human carcinoembryonic
            antigen gene: DNA sequence and homology studies
  JOURNAL   DNA Seq. 4 (3), 185-196 (1993)
  MEDLINE   94214170
   PUBMED   8161821
REFERENCE   3  (bases 1 to 11288)
  AUTHORS   Richards,C.A., Austin,E.A. and Huber,B.E.
  TITLE     Transcriptional regulatory sequences of carcinoembryonic antigen:
            identification and use with cytosine deaminase for tumor-specific
            gene therapy
  JOURNAL   Hum. Gene Ther. 6 (7), 881-893 (1995)
  MEDLINE   96097131
   PUBMED   7578407
REFERENCE   4  (bases 1 to 11288)
  AUTHORS   Richards,C.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-FEB-1993) Richards C. A., Burroughs Wellcome Co.,
            Cell Biology, 3030 Cornwallis Rd, Research Triangle Park, North
            Carolina, USA, 27709
COMMENT     Overlapping sequence: U17131.
FEATURES             Location/Qualifiers
     source          1. .11288
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="lambdaCEA1"
                     /clone_lib="library LL19NL01 from ATCC"
                     /germline
     repeat_region   520. .857
                     /citation=[2]
                     /rpt_family="Alu"
                     /rpt_type=DISPERSED
     repeat_region   2106. .2155
                     /citation=[2]
                     /rpt_family="CA"
                     /rpt_type=TANDEM
                     /rpt_unit=2106. .2107
```

```
     repeat_region   2308. .2501
                     /citation=[2]
                     /rpt_type=TANDEM
                     /rpt_unit=2344. .2379
     repeat_region   6746. .6887
                     /citation=[2]
                     /rpt_family="Alu"
                     /rpt_type=DISPERSED
     mRNA            10694. .11288
                     /citation=[1]
     exon            10694. .10867
                     /citation=[1]
     CDS             10804. .10938
                     /citation=[1]
                     /codon_start=1
                     /product="carcinoembryonic antigen"
                     /protein_id="CAA79884.1"
                     /db_xref="GI:437784"
                     /db_xref="SPTREMBL:Q14081"
                     /translation="MESPSAPPHRWCIPWQRLLLTGEGRTTWERVGGGSWGLLGRTGL
                     "
     intron          10868. .11288
                     /citation=[1]
BASE COUNT     2944 a   3060 c   2955 g   2329 t
ORIGIN
```

```
  Query Match             99.8%;  Score 471;  DB 9;  Length 11288;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-141;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10294 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 10353

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10354 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 10413

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10414 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 10473

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10474 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 10533

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10534 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 10593

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10594 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 10653

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10654 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 10713

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10714 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 10764
```

```
RESULT 6
AR266427
LOCUS       AR266427               15056 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 10 from patent US 6495130.
ACCESSION   AR266427
VERSION     AR266427.1  GI:29695383
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
```

```
REFERENCE   1  (bases 1 to 15056)
  AUTHORS   Henderson,D.R. and Yu,D.C.
  TITLE     Target cell-specific adenoviral vectors containing E3 and methods
            of use thereof
  JOURNAL   Patent: US 6495130-A 10 17-DEC-2002;
FEATURES             Location/Qualifiers
     source          1. .15056
                     /organism="unknown"
BASE COUNT     3850 a   4131 c   4003 g   3072 t
ORIGIN
```

```
  Query Match             99.8%;  Score 471;  DB 6;  Length 15056;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-141;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532
```

```
RESULT 7
AX259954
LOCUS       AX259954               15056 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 39 from Patent WO0172341.
ACCESSION   AX259954
VERSION     AX259954.1  GI:16509014
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Yu,D.C., Chen,Y. and Henderson,D.R.
  TITLE     Methods of treating neoplasia with combinations of target
            cell-specific adenovirus, chemotherapy and radiation
  JOURNAL   Patent: WO 0172341-A 39 04-OCT-2001;
            Calydon, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .15056
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="CEA-TRE"
BASE COUNT     3848 a   4125 c   4011 g   3072 t
ORIGIN
```

```
  Query Match             99.8%;  Score 471;  DB 6;  Length 15056;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-141;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532


RESULT 8
AX262359
LOCUS       AX262359               15056 bp    DNA     linear   PAT 26-OCT-2001
DEFINITION  Sequence 14 from Patent WO0173093.
ACCESSION   AX262359
VERSION     AX262359.1  GI:16511287
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Yu,D.C., Li,Y. and Henderson,D.R.
  TITLE     Cell-specific adenovirus vectors comprising an internal ribosome
            entry site
  JOURNAL   Patent: WO 0173093-A 14 04-OCT-2001;
            Calydon, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .15056
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="CEA-TRE"
BASE COUNT     3848 a   4125 c   4011 g   3072 t
ORIGIN

  Query Match             99.8%;  Score 471;  DB 6;  Length 15056;
  Best Local Similarity   100.0%;  Pred. No. 7.5e-141;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532


RESULT 9
AC008999
LOCUS       AC008999               39707 bp    DNA     linear   PRI 23-AUG-2001
DEFINITION  Homo sapiens chromosome 19 clone LLNLR-269B9, complete sequence.
ACCESSION   AC008999
VERSION     AC008999.7  GI:15281189
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 39707)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 39707)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 39707)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 39707)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   5  (bases 1 to 39707)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Aug 23, 2001 this sequence version replaced gi:11079409.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES             Location/Qualifiers
     source          1. .39707
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
```

```
                     /chromosome="19"
                     /clone="LLNLR-269B9"
BASE COUNT     11027 a    9538 c    8739 g   10403 t
ORIGIN

  Query Match             99.8%;  Score 471;  DB 9;  Length 39707;
  Best Local Similarity   100.0%;  Pred. No. 8.4e-141;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3492 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 3551

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3552 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 3611

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3612 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 3671

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3672 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 3731

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3732 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 3791

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3792 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 3851

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3852 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 3911

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3912 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 3962


RESULT 10
HSCEATG
LOCUS       HSCEATG                 4137 bp    DNA     linear   PRI 07-MAR-1997
DEFINITION  H.sapiens carcinoembryonic antigen gene.
ACCESSION   X62151
VERSION     X62151.1  GI:29856
KEYWORDS    carcinoembryonic antigen.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4137)
  AUTHORS   Barnett,T.
  TITLE     Genomic DNA sequence upstream of the translational start of the
            carcinoembryonic antigen gene
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4137)
  AUTHORS   Barnett,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-SEP-1991) T. Barnett, Miles, Inc., 400 Morgan Lane,
            West Haven, CT 06516, USA
FEATURES             Location/Qualifiers
     source          1. .4137
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="CEA4kbN-term"
                     /cell_line="AG7647 (NIGMS Repository)"
                     /cell_type="EBV-immortalized lymphoblast"
                     /clone_lib="EcoRI segment of cosmid; c103.2 from Alzheimer
                     6 cosmid library"
     5'UTR           1. .2616
     CDS             join(2617. .2680,3605. .3972)
                     /codon_start=1
                     /product="carinoembryonic antigen"
                     /protein_id="CAA44076.1"
                     /db_xref="GI:1877203"
                     /db_xref="SPTREMBL:P7B448"
                     /translation="MESPSAPPHRWCIPWQRLLLTASLLTFWNPPTTAKLTIESTPFN
                     VAEGKEVLLLVHNLPQHLFGYSWYKGERVDGNRQIIGYVIGTQQATPGPAYSGREIIY
                     PNASLLIQNIIQNDTGFYTLHVIKSDLVNEEATGQFRVYRE"
     repeat_unit     3469. .3500
                     /rpt_family="(AC)"
BASE COUNT     1048 a    1200 c    1059 g     830 t
ORIGIN

  Query Match             96.6%;  Score 455.8;  DB 9;  Length 4137;
  Best Local Similarity   99.4%;  Pred. No. 5.4e-136;
  Matches  468;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2108 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 2167

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2168 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 2227

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2228 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 2287

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2288 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTGGATCCACCTT 2347

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2348 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 2407

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db       2408 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCC-TGACCCA 2466

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2467 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 2526

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2527 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 2577


RESULT 11
A37261
LOCUS       A37261                  3281 bp    DNA     linear   PAT 05-MAR-1997
DEFINITION  Sequence 1 from Patent WO9404196.
ACCESSION   A37261
VERSION     A37261.1  GI:2294358
KEYWORDS    .
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 3281)
  AUTHORS   Vile,R.G. and Hart,I.R.
  TITLE     TUMOUR THERAPY
  JOURNAL   Patent: WO 9404196-A 1 03-MAR-1994;
            IMP CANCER RES TECH (GB)
FEATURES             Location/Qualifiers
     source          1. .3281
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
BASE COUNT      847 a     953 c     871 g     610 t
```

ORIGIN

Query Match 91.6%; Score 432.4; DB 6; Length 3281;
Best Local Similarity 98.9%; Pred. No. 2e-128;
Matches 467; Conservative 0; Mismatches 1; Indels 4; Gaps 3;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 1320

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1321 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGA-CCTGCTGGG 1379

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 1439

Qy        181 CAGAGGTCAGCACTGGGG-AAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCT 239
              |||||||||||||||||| |||||||||||||  | ||||||||||||||||||||||||
Db       1440 CAGAGGTCAGCACTGGGGAAAGACAGGTTGTC--CACAGGGGATGGGGGTCCATCCACCT 1497

Qy        240 TGCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAG 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 TGCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAG 1557

Qy        300 AGGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCC 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 AGGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCC 1617

Qy        360 ACGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGAC 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 ACGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGAC 1677

Qy        420 AGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 1729
```

RESULT 12
HUMCEA01
LOCUS HUMCEA01 3281 bp DNA linear PRI 06-MAR-1995
DEFINITION Human carcinoembryonic antigen (CEA) gene, complete cds.
ACCESSION M59255 M31966
VERSION M59255.1 GI:180200
KEYWORDS carcinoembryonic antigen.
SEGMENT 1 of 10
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3281)
AUTHORS Schrewe,H., Thompson,J., Bona,M., Hefta,L.J., Maruya,A., Hassauer,M., Shively,J.E., von Kleist,S. and Zimmermann,W.
TITLE Cloning of the complete gene for carcinoembryonic antigen: analysis of its promoter indicates a region conveying cell type-specific expression
JOURNAL Mol. Cell. Biol. 10 (6), 2738-2748 (1990)
MEDLINE 90258861
PUBMED 2342461
COMMENT Original source text: Human DNA.
FEATURES Location/Qualifiers
source 1. .3281
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="19q13.2"
/clone="cosCEA1"
exon 1659. .1832
/gene="CEA"
/note="G00-119-054"
/number=1
sig_peptide join(1769. .1832,2725. .2762)
/gene="CEA"
/note="G00-119-054"
exon 2725. .3084
/gene="CEA"
/note="G00-119-054"
/number=2
BASE COUNT 847 a 953 c 871 g 610 t
ORIGIN

Query Match 91.6%; Score 432.4; DB 9; Length 3281;
Best Local Similarity 98.9%; Pred. No. 2e-128;
Matches 467; Conservative 0; Mismatches 1; Indels 4; Gaps 3;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 1320

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1321 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGA-CCTGCTGGG 1379

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 1439

Qy        181 CAGAGGTCAGCACTGGGG-AAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCT 239
              |||||||||||||||||| |||||||||||||  | ||||||||||||||||||||||||
Db       1440 CAGAGGTCAGCACTGGGGAAAGACAGGTTGTC--CACAGGGGATGGGGGTCCATCCACCT 1497

Qy        240 TGCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAG 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 TGCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAG 1557

Qy        300 AGGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCC 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 AGGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCC 1617

Qy        360 ACGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGAC 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 ACGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGAC 1677

Qy        420 AGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 1729
```

RESULT 13
AX705383
LOCUS AX705383 3500 bp DNA linear PAT 04-APR-2003
DEFINITION Sequence 52 from Patent WO03014388.
ACCESSION AX705383
VERSION AX705383.1 GI:29562048
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Distler,J., Model,F. and Taubert,H.
TITLE Method and nucleic acids for the analysis of colon cancer
JOURNAL Patent: WO 03014388-A 52 20-FEB-2003;
Epigenomics AG (DE)
FEATURES Location/Qualifiers
source 1. .3500
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 843 a 59 c 931 g 1667 t
ORIGIN

Query Match 58.9%; Score 278.2; DB 6; Length 3500;
Best Local Similarity 75.4%; Pred. No. 1.8e-78;
Matches 346; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

```
Qy         12 AGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTGGGCATCATCCC 71
Db       2517 AGTGAGTTTTTTTTTAGTTTTTAGAGTTATTTTTGTTATTTTTTTGTTGGGTATTATTTT 2576

Qy         72 ACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGGTTTCTCTGTCA 131
Db       2577 ATTTTTTTAGAGTTTTGGAGAGTATGGGGAGATTCGGGATTTTGTTGGGTTTTTTTGTTA 2636

Qy        132 CAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAGCAGAGGTCAGC 191
Db       2637 TAAAGGAAAATAATTTTTTTGGTGTGATAGATTTAAGGATAGAATATAGTAGAGGTTAGT 2696

Qy        192 ACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTTGCCGAAAAGAT 251
Db       2697 ATTGGGGAAGATAGGTTGTTTTTTTAGGGGATGGGGGTTTATTTATTTTGTCGAAAAGAT 2756

Qy        252 TTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAAGAGGAGGGACAAAAGAGGCAGAAATGA 311
Db       2757 TTGTTTGAGGAATTGAAAATAGAAGGGAAAAAAGAGGAGGGATAAAAGAGGTAGAAATGA 2816

Qy        312 GAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCACGTGATGCTGA 371
Db       2817 GAGGGGAGGGGATAGAGGATATTTGAATAAAGATTATATTTATGATTTACGTGATGTTGA 2876

Qy        372 GAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAGCAGACCAGAC 431
Db       2877 GAAGTATTTTTGTTTTAGGAAGAGATTTAGGGTAGAGGGAGGAAGGATAGTAGATTAGAT 2936

Qy        432 AGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAG 470
Db       2937 AGTTATAGTAGTTTTGATAAAACGTTTTTGGAATTTAAG 2975
```

```
RESULT 14
AX705405
LOCUS       AX705405                3500 bp    DNA     linear   PAT 04-APR-2003
DEFINITION  Sequence 74 from Patent WO03014388.
ACCESSION   AX705405
VERSION     AX705405.1  GI:29562070
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Distler,J., Model,F. and Taubert,H.
  TITLE     Method and nucleic acids for the analysis of colon cancer
  JOURNAL   Patent: WO 03014388-A 74 20-FEB-2003;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .3500
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT      843 a      0 c    931 g   1726 t
ORIGIN
```

```
  Query Match             57.6%;  Score 271.8;  DB 6;  Length 3500;
  Best Local Similarity   74.5%;  Pred. No. 2.2e-76;
  Matches  342;  Conservative    0;  Mismatches  117;  Indels    0;  Gaps    0;

Qy         12 AGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTGGGCATCATCCC 71
Db       2517 AGTGAGTTTTTTTTTAGTTTTTAGAGTTATTTTTGTTATTTTTTTGTTGGGTATTATTTT 2576

Qy         72 ACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGGTTTCTCTGTCA 131
Db       2577 ATTTTTTTAGAGTTTTGGAGAGTATGGGGAGATTTGGGATTTTGTTGGGTTTTTTTGTTA 2636

Qy        132 CAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAGCAGAGGTCAGC 191
Db       2637 TAAAGGAAAATAATTTTTTTGGTGTGATAGATTTAAGGATAGAATATAGTAGAGGTTAGT 2696
```

```
Qy        192 ACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTTGCCGAAAAGAT 251
Db       2697 ATTGGGGAAGATAGGTTGTTTTTTTAGGGGATGGGGGTTTATTTATTTTGTTGAAAAGAT 2756

Qy        252 TTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAAGAGGAGGGACAAAAGAGGCAGAAATGA 311
Db       2757 TTGTTTGAGGAATTGAAAATAGAAGGGAAAAAAGAGGAGGGATAAAAGAGGTAGAAATGA 2816

Qy        312 GAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCACGTGATGCTGA 371
Db       2817 GAGGGGAGGGGATAGAGGATATTTGAATAAAGATTATATTTATGATTTATGTGATGTTGA 2876

Qy        372 GAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAGCAGACCAGAC 431
Db       2877 GAAGTATTTTTGTTTTAGGAAGAGATTTAGGGTAGAGGGAGGAAGGATAGTAGATTAGAT 2936

Qy        432 AGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAG 470
Db       2937 AGTTATAGTAGTTTTGATAAAATGTTTTTGGAATTTAAG 2975
```

```
RESULT 15
HSCGM2PRO
LOCUS       HSCGM2PRO               645 bp    DNA     linear   PRI 30-JUN-1997
DEFINITION  H.sapiens carcinoembryonic antigen family member 2, CGM2, promoter
            region.
ACCESSION   X98312
VERSION     X98312.1  GI:1524061
KEYWORDS    carcinoembryonic antigen; carcinoembryonic antigen family member 2;
            CGM2 gene.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Thompson,J., Seitz,M., Chastre,E., Ditter,M., Aldrian,C.,
            Gespach,C. and Zimmermann,W.
  TITLE     Down-regulation of carcinoembryonic antigen family member 2
            expression is an early event in colorectal tumorigenesis
  JOURNAL   Cancer Res. 57 (9), 1776-1784 (1997)
  MEDLINE   97280695
   PUBMED   9135022
REFERENCE   2  (bases 1 to 645)
  AUTHORS   Zimmermann,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-1996) W. Zimmermann, Albert-Ludwigs-University,
            Institute of Immunobiology, Stefan-Meier-Strasse 8, D-79104
            Freiburg, FRG
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /map="q13.2"
                     /clone="cosCGM2"
                     /tissue_type="placenta"
                     /dev_stage="embyro"
                     /germline
     gene            643. .645
                     /gene="CGM2"
     CDS             643. .>645
                     /gene="CGM2"
                     /note="family member 2"
                     /codon_start=1
                     /product="carcinoembryonic antigen"
                     /protein_id="CAA66956.1"
                     /db_xref="GI:4378976"
                     /translation="M"
BASE COUNT      165 a    199 c    175 g    106 t
ORIGIN
```

Query Match 56.1%; Score 264.8; DB 9; Length 645;
Best Local Similarity 80.3%; Pred. No. 3.3e-74;
Matches 384; Conservative 0; Mismatches 62; Indels 32; Gaps 5;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              || || | ||||||||||  |||||| ||||||||||||||||||||||||||||||| ||
Db        155 AGTCATCCCCCAGTGAGCCCCTTTCTATCCCCCAGAGCCACCTCTGTCACCTTCCTGCTG 214

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              |||||||||||||                          | || ||||  | ||| ||||||
Db        215 GGCATCATCCCAC-----------------------GAGAAACCCAAGGCCCAGCTGGG 250

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              |||||||||||| |||| ||||||||||||||||||||||||||||||| ||||||||||
Db        251 TTTCTCTGTCAC-AAGGGAAATAATCCCCCTGGTGTGACAGACCCAAGGCCAGAACACAG 309

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              || |||||||||||||| |||   ||||||||||||||||  ||||||||||  ||||
Db        310 CATAGGTCAGCACTGGGAAAGGTTGGTTGTCCTCCCAGGGGACAGGGGTCCATCAGCCTT 369

Qy        241 GC--CGAAAAGATTTGTCTGAGGAACTGAAAATAG----AAGGGAAAAAAGAGGAGGGAC 294
              ||  |  |||| ||||| |  ||||||||| ||||    |||||||| | || |||||||
Db        370 GCTACTCAAAGCTTTGTATCGGGAACTGAATATAGACAAAAGGGAAAGAGGAAGAGGGAC 429

Qy        295 AAAAGAGGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCAT 354
              |||||||||| |||||||||||||||||||||||||| |||||  | ||||| ||||| |
Db        430 AAAAGAGGCAGGAATGAGAGGGGAGGGGACAGAGGACTCCTGAGCAGAGACCCCACCCCT 489

Qy        355 GACCCACGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGA 414
              |  ||| |||||  ||  |||| |||||||||| ||| ||| |||||  |||||| ||||
Db        490 GGACCATGTGATCTTGGAAAGTGCTCCTGCCCTGGGAGGAGGCTCAGCACAGAGGTAGGA 549

Qy        415 AGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAA-AACGTTCCTGGAACTCAAGC 471
              |||||||||| | |||||||||||||||| |||||| |  |||||||| | |||||
Db        550 AGGACAGCAGAGCTGACAGTCACAGCAGCCCTGACAAGAGAGTTCCTGGAGCCCAAGC 607
```

Search completed: September 26, 2003, 13:06:57
Job time : 1910 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 26, 2003, 10:11:26 ; Search time 248 Seconds
(without alignments)
5137.640 Million cell updates/sec

Title: US-10-045-116-1
Perfect score: 472
Sequence: 1 AGCCACCACCCAGTGAGCCT.....ACGTTCCTGGAACTCAAGCA 472

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	472	100.0	472	19	AAV52944	CEA transcriptiona
	2	471	99.8	11288	16	AAQ90512	CEA clone HindIII-
	3	471	99.8	11288	23	AAS14778	Human carcinoembry
	4	471	99.8	15056	19	AAV52967	Carcinoembryonic a
	5	471	99.8	15056	21	AAA46851	Nucleotide sequenc
	6	471	99.8	15056	21	AAZ99933	DNA sequence of co
	7	471	99.8	15056	22	AAH43620	CEA-TRE. Homo sap
	8	471	99.8	15056	23	AAF87238	CEA-TRE fusion pro
	9	471	99.8	15056	24	ABK99582	Prostate-specific
	10	359	76.1	14556	16	AAQ90511	CEA genomic clone.
	11	225	47.7	271	20	AAZ09584	Carcinoembryonic a
c	12	120.8	25.6	39801	24	ABK83462	Human cDNA differe
	13	94	19.9	3774	16	AAQ73990	CEA clone HindIII
	14	85.8	18.2	39801	24	ABK83462	Human cDNA differe
	15	75	15.9	2974	24	ABN95819	Gene #2317 used to
	16	75	15.9	2974	24	ABL64746	Lung cancer relate
	17	75	15.9	2974	25	ABX76144	Lung cancer-associ
	18	75	15.9	2974	25	ABX76396	Lung cancer-associ
	19	75	15.9	2974	25	ABQ83855	Human CEA encoding
	20	75	15.9	2975	24	ABQ82535	Human CEA nucleoti
	21	69.4	14.7	653	24	ABQ60586	Human colon cancer
	22	60.6	12.8	474	22	ABA59345	Human foetal liver
	23	60.6	12.8	474	22	ABA28046	Probe #6512 for ge
	24	60.6	12.8	474	22	AAK07563	Human brain expres
	25	60.6	12.8	474	22	AAK33359	Human bone marrow
	26	60.6	12.8	474	22	AAI16427	Probe #6360 for ge
	27	60.6	12.8	474	22	AAI39146	Probe #7832 used t
	28	60.6	12.8	474	23	ABS33113	Human liver single
	29	60.6	12.8	474	24	ABS08194	Human genome-deriv
	30	59	12.5	1298	21	AAC77896	Human cancer assoc
	31	57.6	12.2	387	20	AAZ08319	Human lung tumour
	32	57.6	12.2	387	21	AAC79171	Human lung tumour-
	33	57.6	12.2	387	23	AAD23246	Human lung tumour-
	34	57.2	12.1	460	21	AAA77889	cDNA encoding huma
	35	57.2	12.1	460	22	AAI28627	Colon tumour relat
	36	57.2	12.1	460	25	ABZ32813	Human colon tumour
	37	56.4	11.9	319	21	AAC10830	Human secreted pro
c	38	55.4	11.7	294	24	ABK44760	cDNA encoding colo
	39	54.8	11.6	306	21	AAC10828	Human secreted pro
	40	54.8	11.6	327	21	AAC10827	Human secreted pro
	41	54.8	11.6	414	21	AAC10829	Human secreted pro
	42	54.6	11.6	504	21	AAC10831	Human secreted pro
	43	54.2	11.5	502	21	AAC04586	Human secreted pro
	44	54.2	11.5	2115	21	AAC10826	Human secreted pro
	45	54	11.4	2220	17	AAT33302	Carcinoembryonic a

ALIGNMENTS

RESULT 1
AAV52944
ID AAV52944 standard; DNA; 472 BP.
XX
AC AAV52944;
XX
DT 21-DEC-1998 (first entry)
XX
DE CEA transcriptional regulatory element (CEA-TRE).
XX
KW Carcinoembryonic antigen; transcriptional regulatory element;
KW CEA-TRE; human; promoter; enhancer; adenovirus; vector; cancer;
KW gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO9839467-A2.
XX
PD 11-SEP-1998.
XX
PF 03-MAR-1998; 98WO-US04133.
XX
PR 02-MAR-1998; 98US-0039763.
PR 03-MAR-1997; 97US-0039763.
XX
PA (CALY-) CALYDON INC.
XX
PI Henerson DR, Lamparski HG, Schuur ER;
XX
DR WPI; 1998-495862/42.
XX

PT New adenovirus vectors, particularly for cancer therapy - comprising
PT adenovirus gene under transcriptional control of carcinoembryonic
PT antigen transcriptional regulatory element
XX
PS Claim 13; Page 62-63; 95pp; English.
XX
CC This 472 nucleotide fragment comprises nucleotides -402 to +69
CC of the human carcinoembryonic antigen transcriptional regulatory
CC element (CEA-TRE). It was isolated from human genomic DNA by PCR
CC (see AAV52945-46). The CEA-TRE is capable of mediating gene
CC expression specific to cells capable of expressing CEA or capable
CC of CRE-TRE-mediated transcription. A claimed replication-competent
CC adenovirus (Ad) vector comprises an Ad gene under transcriptional
CC control of a CEA-TRE. Also claimed are: (1) a host cell transformed
CC with a Ad vector as above; (2) a method of making a masked Ad, and
CC (3) an Ad complexed with a masking agent. By providing for
CC transcriptional initiating regulation dependent upon CEA
CC expression, virus replication can be restricted to target cells
CC which allow a CEA-TRE to function, particularly carcinoma cells
CC expressing CEA. The vectors can be used to detect and monitor
CC samples for the presence of cells that allow a CEA-TRE to function,
CC and to selectively kill such cells, especially malignant cells.
CC Preferred vectors contain a CEA-TRE comprising nucleotides 313-472
CC or 104-472 of the 472 nucleotide fragment, especially comprising an
CC enhancer and/or a promoter of the CEA gene.
XX
SQ Sequence 472 BP; 142 A; 125 C; 134 G; 71 T; 0 other;

Query Match 100.0%; Score 472; DB 19; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.1e-127;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
```

RESULT 2
AAQ90512
ID AAQ90512 standard; DNA; 11288 BP.
XX
AC AAQ90512;
XX
DT 25-MAR-2003 (updated)
DT 01-NOV-1995 (first entry)
XX
DE CEA clone HindIII-Sau3A fragment.
XX
KW Carcinoembryonic antigen; CEA; transcription regulatory sequence;
KW TRS; gene targeting; cancer; metastasis; gene therapy;
KW cytosine deaminase; ss.
XX
OS Homo sapiens.
XX
PN WO9514100-A2.
XX
PD 26-MAY-1995.
XX
PF 18-NOV-1994; 94WO-GB02546.
XX
PR 19-NOV-1993; 93US-0154712.
XX
PA (WELL) WELLCOME FOUND LTD.
XX
PI Huber B, Richards CA;
XX
DR WPI; 1995-200389/26.
XX
PT New carcinoembyronic antigen transcriptional regulatory sequence DNA -
PT used partic. for expressing heterologous enzymes for pro-drugs in the
PT treatment of cancers
XX
PS Disclosure; Page 31-37; 64pp; English.
XX
CC CEA genomic clone lambdaCEA1 was isolated from human chromosome 19
CC genomic library LL19NL01 (ATCC 57766). An HindIII/Sau3A fragment
CC of the clone extended from -10.7 to +0.6 kb relative to the start
CC site of CEA mRNA. TRS regions of CEA are used to target e.g.
CC cytosine deaminase to cancer cells for prodrug activation.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 11288 BP; 2940 A; 3063 C; 2953 G; 2332 T; 0 other;

Query Match 99.8%; Score 471; DB 16; Length 11288;
Best Local Similarity 100.0%; Pred. No. 8.2e-127;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10294 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 10353

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10354 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 10413

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10414 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 10473

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10474 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 10533

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10534 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 10593

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10594 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 10653

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10654 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 10713

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10714 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 10764
```

```
RESULT 3
AAS14778
ID   AAS14778 standard; DNA; 11288 BP.
XX
AC   AAS14778;
XX
DT   27-FEB-2002  (first entry)
XX
DE   Human carcinoembryonic antigen (CEA) genomic DNA.
XX
KW   Recombinant nucleic acid vector; carcinoembryonic antigen; CEA; cytokine;
KW   syncytium-inducing polypeptide; fusogenic membrane glycoprotein; tumour;
KW   recombinase; tumour-specific promoter; hypoxic response element; HRE; ds;
KW   tyrosinase promoter; Cre; FLP; retroviral vector; malignant cell; cancer;
KW   cytostatic; gene therapy; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             10804..10938
FT                   /*tag=  a
FT                   /product= "Human carcinoembryonic antigen"
XX
PN   WO200174861-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US10250.
XX
PR   31-MAR-2000; 2000US-193977P.
XX
PA   (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.
XX
PI   Vile RG,  Harrington K,  Murphy S,  Bateman A;
XX
DR   WPI; 2001-656985/75.
XX
PT   Recombinant nucleic acid vector for reducing tumour size, has expression
PT   cassette comprises a promoter linked to nucleic acid sequence encoding
PT   a syncytium-inducing polypeptide and flanked on either side by
PT   recombinase -
XX
PS   Disclosure; Fig 3; 84pp; English.
XX
CC   The invention relates to a recombinant nucleic acid vector comprising a
CC   first expression cassette, comprising a first promoter operably linked to
CC   a nucleic acid sequence encoding a syncytium-inducing polypeptide (such
CC   as a fusogenic membrane glycoprotein) and flanked on either side by a
CC   sequence recognised by a recombinase, and/or a second expression cassette
CC   comprising a tumour-specific promoter operably linked to a nucleic acid
CC   sequence encoding a recombinase. The nucleic acid of the first expression
CC   cassette may be linked to a hypoxic response element (HRE), the second
CC   expression cassette may contain a promoter linked to a nucleic acid
CC   encoding a cytokine, and a third cassette may contain a tumour specific
CC   promoter linked to the nucleic acid encoding the recombinase. The tumour
CC   specific promoter is, for example, a carcinoembryonic antigen (CEA)
CC   promoter or a tyrosinase promoter and the recombinase is, for example,
CC   Cre recombinase or FLP recombinase. The invention is useful for reducing
CC   tumour size by administering the compositions as retroviral vectors, or
CC   in a cell containing the vector, to an individual in need of treatment
CC   for a disease caused by malignant cells. This sequence represents genomic
CC   DNA encoding the human carcinoembryonic antigen (CEA).
XX
SQ   Sequence 11288 BP; 2944 A; 3060 C; 2955 G; 2329 T; 0 other;

  Query Match             99.8%;  Score 471;  DB 23;  Length 11288;
  Best Local Similarity   100.0%;  Pred. No. 8.2e-127;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      10294 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 10353

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      10354 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 10413

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      10414 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 10473

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      10474 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 10533

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      10534 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 10593

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      10594 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 10653

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db      10654 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 10713

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||

Db      10714 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 10764

RESULT 4
AAV52967
ID   AAV52967 standard; DNA; 15056 BP.
XX
AC   AAV52967;
XX
DT   21-DEC-1998  (first entry)
XX
DE   Carcinoembryonic antigen gene 5' flanking region.
XX
KW   Carcinoembryonic antigen; transcriptional regulatory element;
KW   CEA-TRE; human; promoter; enhancer; adenovirus; vector; cancer;
KW   gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    14062..14533
FT                   /*tag=  a
FT                   /label= CEA-TRE
FT                   /note= "nucleotides -402 to +69 (Claim 13)"
XX
PN   WO9839467-A2.
XX
PD   11-SEP-1998.
XX
PF   03-MAR-1998;   98WO-US04133.
XX
PR   02-MAR-1998;   98US-0039763.
PR   03-MAR-1997;   97US-0039763.
XX
PA   (CALY-) CALYDON INC.
XX
PI   Henerson DR,  Lamparski HG,  Schuur ER;
XX
DR   WPI; 1998-495862/42.
XX
PT   New adenovirus vectors, particularly for cancer therapy - comprising
PT   adenovirus gene under transcriptional control of carcinoembryonic
PT   antigen transcriptional regulatory element
XX
PS   Disclosure; Fig 2A-K; 95pp; English.
```

XX
CC This nucleotide sequence comprises the 5' flanking region of the
CC human carcinoembryonic antigen (CEA) gene to nucleotide +537
CC relative to the transcrpitional start. A 472 bp fragment (see
CC AAV52944) of this 5' flanking region (nucleotides -402 to +69)
CC comprises the CEA transcriptional regulatory element (CEA-TRE) that
CC is capable of mediating gene expression specific to cells capable
CC of expressing CEA or capable of CRE-TRE-mediated transcription. A
CC claimed replication-competent adenovirus (Ad) vector comprises an
CC Ad gene under transcriptional control of a CEA-TRE. By providing
CC for transcriptional initiating regulation dependent upon CEA
CC expression, virus replication can be restricted to target cells
CC which allow a CEA-TRE to function, particularly carcinoma cells
CC expressing CEA. The vectors can be used to detect and monitor
CC samples for the presence of cells that allow a CEA-TRE to function,
CC and to selectively kill such cells, especially malignant cells.
XX
SQ Sequence 15056 BP; 3848 A; 4125 C; 4011 G; 3072 T; 0 other;

Query Match 99.8%; Score 471; DB 19; Length 15056;
Best Local Similarity 100.0%; Pred. No. 9.2e-127;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532
```

RESULT 5
AAA46851
ID AAA46851 standard; DNA; 15056 BP.
XX
AC AAA46851;
XX
DT 03-OCT-2000 (first entry)
XX
DE Nucleotide sequence of a CEA-TRE.
XX
KW Adenoviral vector; adenoviral gene; E3 sequence; cancer;
KW target cell-specific transcriptional regulatory element; TRE;
KW selective cytotoxicity; cell growth; tumour growth; ss.
XX
OS Homo sapiens.
XX
PN WO200039319-A2.

XX
PD 06-JUL-2000.
XX
PF 30-DEC-1999; 99WO-US31249.
XX
PR 30-DEC-1998; 98US-0114262.
PR 29-DEC-1999; 99US-0474699.
XX
PA (CALY-) CALYDON INC.
XX
PI Henderson DR, Yu D;
XX
DR WPI; 2000-452409/39.
XX
PT New adenoviral vector for selective targeting of cancer cells,
PT comprises an adenovirus gene under transcriptional control of a target
PT cell-specific transcriptional regulatory element and an E3 sequence -
XX
PS Disclosure; Fig 23; 138pp; English.
XX
CC The specification describes an adenoviral vector which comprises an
CC adenoviral gene under the transcriptional control of a target
CC cell-specific transcriptional regulatory element (TRE) and an E3
CC sequence. The vectors are useful for exerting selective cytotoxicity
CC and/or suppressing cell growth (including tumour growth). This is
CC especially useful in treating cancer in which targeted cell killing is
CC desirable. This is also useful for targeted cytotoxic effects in other,
CC non-tumour cells, when selective destruction and/or suppression of these
CC cells is desirable. The vectors can also be useful for detecting the
CC presence of cells which permit function of a target cell-specific TRE in,
CC for example, an appropriate biological (such as clinical) sample. The
CC adenovirus vector(s) can optionally selectively produce one or more
CC proteins of interest in a target cell by using a heterologous, i.e.,
CC non-adenoviral, TRE. The present sequence represents a CEA-TRE, which
CC is used in the course of the invention.
XX
SQ Sequence 15056 BP; 3848 A; 4125 C; 4011 G; 3072 T; 0 other;

Query Match 99.8%; Score 471; DB 21; Length 15056;
Best Local Similarity 100.0%; Pred. No. 9.2e-127;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532
```

```
RESULT 6
AAZ99933
ID   AAZ99933 standard; DNA; 15056 BP.
XX
AC   AAZ99933;
XX
DT   25-JUL-2000  (first entry)
XX
DE   DNA sequence of comprising a carcinoembryonic antigen TRE.
XX
KW   Carcinoembryonic antigen; adenoviral vector; adenovirus gene;
KW   transcriptional regulatory element; TRE; transcriptional control;
KW   adenoviral propagation; tumour; ss.
XX
OS   Unidentified.
XX
PN   WO200015820-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999;   99WO-US20718.
XX
PR   10-SEP-1998;   98US-0099791.
PR   09-SEP-1999;   99US-0099791.
XX
PA   (CALY-) CALYDON INC.
XX
PI   Yu DC,  Henderson DR;
XX
DR   WPI; 2000-271456/23.
XX
PT   Adenovirus vectors comprising cell-status specific response elements
PT   useful in gene therapy protocols for the treatment of cancers -
XX
PS   Disclosure; Fig 5A-I; 79pp; English.
XX
CC   The present sequence comprises a transcriptional regulatory element
CC   (TRE) from a carcinoembryonic antigen gene. The TRE is used to
CC   produce an adenoviral vector of the invention. The specification
CC   describes an adenovirus vector which comprises an adenovirus gene
CC   under transcriptional control of a cell status specific TRE. The
CC   TRE is preferably one that is essential for adenoviral propagation.
CC   The adenovirus vectors may be used for the treatment of a range of
CC   tumours such as lung, stomach, breast, colon and rectum, and uterine
CC   and cervix cancers.
XX
SQ   Sequence 15056 BP; 3848 A; 4125 C; 4011 G; 3072 T; 0 other;

  Query Match             99.8%;  Score 471;  DB 21;  Length 15056;
  Best Local Similarity   100.0%;  Pred. No. 9.2e-127;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532


RESULT 7
AAH43620
ID   AAH43620 standard; cDNA; 15056 BP.
XX
AC   AAH43620;
XX
DT   07-JAN-2002  (first entry)
XX
DE   CEA-TRE.
XX
KW   Adenovirus; ADP; replication-competent; adenoviral vector; TRE; PCR;
KW   transcriptional regulatory element; mutation; deletion; IRES; primer;
KW   promoter; internal ribosome entry site; cytotoxic; cancer; bladder;
KW   amplify; polymerase chain reaction; alpha-fetoprotein; AFP; EMCV;
KW   encephalomycarditis virus; vascular endothelial growth factor; VEGF;
KW   immunoglobulin heavy-chain binding protein; BiP; uroplakin II; PDGF;
KW   platelet derived growth factor; hypoxia responsive element; HRE;
KW   prostate-specific antigen; PSA; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   precursor_RNA   14466..15056
FT                   /*tag=  a
FT                   /note= "Transcription start site"
XX
PN   WO200173093-A2.
XX
PD   04-OCT-2001.
XX
PF   21-MAR-2001; 2001WO-US09036.
XX
PR   24-MAR-2000; 2000US-192156P.
XX
PA   (CALY-) CALYDON INC.
XX
PI   Yu D,  Li Y,  Henderson DR;
XX
DR   WPI; 2001-639234/73.
XX
PT   Replication-competent adenoviral vector, useful e.g. for killing cancer
PT   cells, contains two genes linked by internal ribosome entry site and
PT   controlled by target-specific regulator -
XX
PS   Example; Page 117-125; 148pp; English.
XX
CC   The sequences given in AAH43607-22 represent sequences which may used
CC   in the replication-competent adenoviral vector (A) of the invention.
CC   The vector contains two genes (G1, G2) that are co-transcribed as a
CC   single mRNA and under control of a heterologous, target cell-specific
CC   transcriptional regulatory element (TRE). G2 has a mutation in, or
CC   deletion of, its endogenous promoter and is controlled from an internal
CC   ribosome entry site (IRES). (A) has greater specificity for a target
CC   cell than a similar vector in which TRE is operably linked to a gene
CC   and which lacks an IRES. (A) are used to modify the genotype of target
CC   cells, optionally in vitro with subsequent return of altered cells to
CC   the host and where G2 is a cytotoxic gene, to confer selective cyto-
CC   toxicity to target cells, especially for killing cancer cells. Also
CC   (A) are used for diagnosis and monitoring, e.g. detection of bladder
```

CC cancer cells. The target cell-specific TRE ensures that (A) has better
CC targeting specificity, with minimal replication in non-target cells, so
CC a runaway infection is prevented but production of adenoviral proteins
CC in target cells activates and/or stimulates the immune response against
CC target cells producing such proteins. The use of an IRES (rather than
CC two identical control elements) eliminates the risk of homologous
CC recombination and may provide enough extra space for an additional
CC (therapeutic) gene.
XX
SQ Sequence 15056 BP; 3848 A; 4125 C; 4011 G; 3072 T; 0 other;

Query Match 99.8%; Score 471; DB 22; Length 15056;
Best Local Similarity 100.0%; Pred. No. 9.2e-127;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532
```

RESULT 8
AAF87238
ID AAF87238 standard; DNA; 15056 BP.
XX
AC AAF87238;
XX
DT 26-MAR-2002 (first entry)
XX
DE CEA-TRE fusion protein coding sequence.
XX
KW Tumour growth suppression; adenovirus vector; antineoplastic agent;
KW transcriptional regulatory element; TRE; radiotherapy; bladder cancer;
KW prostate cancer; liver cancer; breast cancer; colon cancer; melanoma;
KW ovarian cancer; pancreatic cancer; lung cancer; brain cancer; therapy;
KW CEA-TRE; ds.
XX
OS Homo sapiens.
XX
PN WO200172341-A2.
XX
PD 04-OCT-2001.
XX
PF 21-MAR-2001; 2001WO-US09042.
XX
PR 24-MAR-2000; 2000US-192015P.
XX
PA (CALY-) CALYDON INC.
XX
PI Yu D, Chen Y, Henderson DR;
XX
DR WPI; 2001-648426/74.
XX
PT Suppression of tumour growth, e.g. liver, bladder or breast cancer,
PT comprises using a synergistic combination of adenovirus vector and
PT antineoplastic agent or radiotherapy -
XX
PS Claim 12; Page 185-193; 248pp; English.
XX
CC This sequence represents a CEA-TRE fusion protein coding sequence
CC used in the scope of the invention. The invention relates to a method
CC for the suppression of tumour growth comprising the administration of:
CC (a) a target cell-specific adenovirus vector comprising an adenoviral
CC gene essential for replication under transcriptional control of a
CC target cell-specific transcriptional regulatory element (TRE); and
CC (b) at least one antineoplastic agent; or (c) a course of radiotherapy
CC where the amount of (a) and/or (b) or (c) is lower than that known to be
CC effective for suppressing tumour growth when administered alone.
CC The method is used for suppression of tumour growth for treatment of
CC e.g. bladder cancer, prostate cancer, liver cancer, breast cancer, colon
CC cancer, melanoma, ovarian cancer, pancreatic cancer, lung cancer or brain
CC cancer. The combinations enhance the efficacy of treatment, allowing
CC lower doses to be administered, reducing toxicity and suffering of the
CC patient.
XX
SQ Sequence 15056 BP; 3848 A; 4126 C; 4010 G; 3072 T; 0 other;

Query Match 99.8%; Score 471; DB 23; Length 15056;
Best Local Similarity 100.0%; Pred. No. 9.2e-127;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532
```

RESULT 9
ABK99582
ID ABK99582 standard; DNA; 15056 BP.
XX
AC ABK99582;

```
XX
DT   21-OCT-2002  (first entry)
XX
DE   Prostate-specific antigen transcriptional regulatory elelment (TRE).
XX
KW   Adenovirus; transcriptional regulatory element; TRE; prostate; liver;
KW   breast cancer; colon cancer; antitumour; gene; ds; probasin; PB; human;
KW   rat; carcinoembryonic antigen; prostate-specific antigen; ADP;
KW   adenovirus death protein; glandular kallikrein.
XX
OS   Unidentified.
XX
PN   US2002068049-A1.
XX
PD   06-JUN-2002.
XX
PF   06-DEC-2000; 2000US-0732169.
XX
PR   10-SEP-1998;   98US-0151376.
XX
PA   (HEND/) HENDERSON D R.
PA   (SCHU/) SCHUUR E R.
XX
PI   Henderson DR,  Schuur ER;
XX
DR   WPI; 2002-582468/62.
XX
PT   Novel adenovirus vector comprises adenovirus gene under transcriptional
PT   control of cell-type specific transcriptional response element for
PT   conferring selective toxicity on target cell and for suppressing tumor
PT   growth -
XX
PS   Disclosure; Fig 15; 83pp; English.
XX
CC   The invention relates to an adenovirus vector (AV) comprising an AV gene
CC   under transcriptional control of a cell type-specific transcriptional
CC   regulatory element (TRE) and optionally a first AV gene under control of
CC   a first cell type-specific TRE and a second gene under control of a
CC   second cell type-specific TRE, where the first and second cell
CC   type-specific TREs are substantially identical. When the vector is
CC   introduced into a cell (e.g. prostate cell, liver cell, breast cancer
CC   cell or colon cancer cell) it allows the cell type-specific TRE to
CC   function, resulting in cytotoxicity. The vector is useful for suppressing
CC   tumour growth of a target cell. This sequence represents a polynucleotide
CC   used in the scope of the invention.
XX
SQ   Sequence 15056 BP; 3848 A; 4125 C; 4011 G; 3072 T; 0 other;

  Query Match           99.8%;  Score 471;  DB 24;  Length 15056;
  Best Local Similarity  100.0%;  Pred. No. 9.2e-127;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532


RESULT 10
AAQ90511
ID   AAQ90511 standard; DNA; 14556 BP.
XX
AC   AAQ90511;
XX
DT   25-MAR-2003  (updated)
DT   01-NOV-1995  (first entry)
XX
DE   CEA genomic clone.
XX
KW   Carcinoembryonic antigen; CEA; transcription regulatory sequence;
KW   TRS; gene targeting; cancer; metastasis; gene therapy;
KW   cytosine deaminase; hybridization; probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514100-A2.
XX
PD   26-MAY-1995.
XX
PF   18-NOV-1994;   94WO-GB02546.
XX
PR   19-NOV-1993;   93US-0154712.
XX
PA   (WELL ) WELLCOME FOUND LTD.
XX
PI   Huber B,  Richards CA;
XX
DR   WPI; 1995-200389/26.
XX
PT   New carcinoembyronic antigen transcriptional regulatory sequence DNA -
PT   used partic. for expressing heterologous enzymes for pro-drugs in the
PT   treatment of cancers
XX
PS   Disclosure; Fig. 6; 64pp; English.
XX
CC   CEA genomic clone lambdaCEA1 was isolated from human chromosome 19
CC   genomic library LL19NL01 (ATCC 57766).  HindIII/Sau3A and HindIII
CC   fragments of the clone were sequenced and combined to obtain the
CC   sequence given in AAQ90511, which extends from -14.4 to +0.6 kb
CC   relative to the start site of CEA mRNA.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 14556 BP; 3721 A; 3995 C; 3874 G; 2966 T; 0 other;

  Query Match           76.1%;  Score 359;  DB 16;  Length 14556;
  Best Local Similarity  96.9%;  Pred. No. 4.3e-94;
  Matches  443;  Conservative    0;  Mismatches    0;  Indels   14;  Gaps    7;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db      13594 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAG--ACCTCTGTCACCTTCCTGTTG 13651

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db      13652 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCAT--GGAGACCCGGGACCCTGCTGGG 13709

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db      13710 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTG--AGACCCAAGGACAGAACACAG 13767
```

```
Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db      13768 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCA--GGATGGGGGTCCATCCACCTT 13825

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db      13826 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGG--AAAAAGAGGAGGGACAAAAGA 13883

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db      13884 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGA--AAAGACCACACCCATGACCCA 13941

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db      13942 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGA--CAGGGCAGAGGGAGGAAGGACA 13999

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTT 457
              |||||||||||||||||||||||||||||||||||||
Db      14000 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTT 14036
```

RESULT 11
AAZ09584
ID AAZ09584 standard; DNA; 271 BP.
XX
AC AAZ09584;
XX
DT 05-NOV-1999 (first entry)
XX
DE Carcinoembryonic antigen homologue promoter.
XX
KW Carcinoembryonic antigen; CEA; promoter; colon cancer; gene therapy;
KW tumour; secretion; tumour cell; ss.
XX
OS Unidentified.
XX
PN CN1216321-A.
XX
PD 12-MAY-1999.
XX
PF 31-OCT-1997; 97CN-0114280.
XX
PR 31-OCT-1997; 97CN-0114280.
XX
PA (SCRE-) SCI RES & TRAINING SECTION MEDICAL MATTE.
XX
PI Jiang Y, Wang X, Xie B;
XX
DR WPI; 1999-431041/37.
XX
PT Cloning carcinoembryonic antigen promoter DNA - used for tumour gene
PT therapy
XX
PS Claim 3; Page 3; 10pp; Chinese.
XX
CC This invention describes a novel DNA sequence which is dissimilar from
CC the carcinoembryonic antigen (CEA) promoter of colon cancer. In CEA
CC secretary tumour gene therapy, the gene segment is recombined with a
CC therapeutic gene, so that killing of tumour cells may be effected
CC selectively while keeping normal cells unaffected. This sequence
CC represents promoter used to describe the method of the invention.
XX
SQ Sequence 271 BP; 100 A; 57 C; 82 G; 32 T; 0 other;

Query Match 47.7%; Score 225; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.1e-55;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        247 AAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGA 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGA 60

Qy        307 AATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCACGTGAT 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCACGTGAT 120

Qy        367 GCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAGCAGAC 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAGCAGAC 180

Qy        427 CAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 225
```

RESULT 12
ABK83462/c
ID ABK83462 standard; cDNA; 39801 BP.
XX
AC ABK83462;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #33.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US30821.
XX
PR 03-OCT-2000; 2000US-237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity -
XX
PS Claim 1; SEQ ID No 33; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having

```
CC   inflammation with an agent that modulates the expression of gene(s)
CC   from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC   modulating GA; M3 is useful for screening an agent capable of modulating
CC   GCA preferably in an inflammation in a tissue; M4 is useful for
CC   detecting an inflammation (especially chronic) in a tissue, an allergic
CC   response in a subject, exposure of a subject to a pathogen or sterile
CC   inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC   glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC   reperfusion injury, ARDS, adult respiratory distress syndrome,
CC   inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC   periodontal disease; also bacterial infection, viral infection,
CC   parasitic infection, protozoal infection, fungal infection and M5 is
CC   useful for treating one of the above conditions. The present
CC   sequence represents a gene differentially expressed in granulocytes.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 39801 BP; 9969 A; 9367 C; 9531 G; 10934 T; 0 other;

  Query Match             25.6%;  Score 120.8;  DB 24;  Length 39801;
  Best Local Similarity   66.1%;  Pred. No. 2.1e-24;
  Matches  292;  Conservative    0;  Mismatches  122;  Indels   28;  Gaps    7;

Qy         18 CCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTGGGCATCATCCCACCTTC 77
              || ||||||| ||| ||||||||||||||||||||||||||| |||| |||||| |||  ||
Db      37616 CCATTTTCTATCCCACAGAGCCACCTCTGTCACCTTCCTGCTGGGTATCATCTCACACTC 37557

Qy         78 CCAGAGCCCTGGAGAGCATGGGGAGACC-CGGGACCCTGCTGGGTTTCTCTGTCACAAAG 136
              || |||   ||| ||||||| |||||||  ||| ||| ||||||| ||| |||||||||
Db      37556 CCTGAGTATTGGGGAGCATGAGGAGACCTGGGGGCCCAGCTGGGTCTCTGTGTCACAAAA 37497

Qy        137 GAAAATAATCCCCCTGGTGTGACAGACCCAAGGACA----------GAACACAGCA 182
              | ||| | | |||| || ||  |||||| ||   |            ||     ||
Db      37496 GGAAACAGTTCCCCAAGTTTGGGAGACCCCAGAGTACCTCTGTTTGTGGTGACATTCCCA 37437

Qy        183 GAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGG--TCCATCCACCTT 240
              |||         | |  |||| ||   |||| | |||||  |||| |||| | ||| |
Db      37436 AAGGGTCAGTGCAGAGGTGACAAGTCACCCTCTCTGGGGACAGGGGACTCCACCAACCCT 37377

Qy        241 GCCGAAAAGATTTGTCTGAGGA------ACTGAAAATAGAAGGGAAAAAAG-AGGAGGGA 293
              ||     | | |    | ||||      | || | | ||||| ||||   | ||||||||
Db      37376 GCTTCTCAAAGTGTGGTTAGGAAACTGTAATGTACACAGAAGAGAAAGGGGAAGGAGGGA 37317

Qy        294 CAAAAGAGGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCA 353
              ||||| ||||||||||||||||||||||     ||    |     ||   || | ||
Db      37316 CAAAAAAGGCAGAAATGAGAGGGGAGGGGCAGAGGGGTGACCTGGGAAGAGCCCCGCCTC 37257

Qy        354 TGACCCACGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGG 413
              || |||   ||   ||| ||||| || ||||||  ||| ||| |||||  |||| |||||
Db      37256 TGCCCC---TGGCCCTGGGAAGTGCTTCTGCCC-GGGAGGAGGCTCAGCACAGAAGGAGG 37201

Qy        414 AAGGACAGCAGACCAGACAGTC 435
              |||| |||||| || ||||| |
Db      37200 AAGGTCAGCAGCCCCGACAGCC 37179


RESULT 13
AAQ73990
ID   AAQ73990 standard; DNA; 3774 BP.
XX
AC   AAQ73990;
XX
DT   25-MAR-2003  (updated)
DT   01-NOV-1995  (first entry)
XX
DE   CEA clone HindIII fragment.
XX
KW   Carcinoembryonic antigen; CEA; transcription regulatory sequence;
KW   TRS; gene targeting; cancer; metastasis; gene therapy;
KW   cytosine deaminase; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514100-A2.
XX
PD   26-MAY-1995.
XX
PF   18-NOV-1994;   94WO-GB02546.
XX
PR   19-NOV-1993;   93US-0154712.
XX
PA   (WELL ) WELLCOME FOUND LTD.
XX
PI   Huber B,  Richards CA;
XX
DR   WPI; 1995-200389/26.
XX
PT   New carcinoembyronic antigen transcriptional regulatory sequence DNA -
PT   used partic. for expressing heterologous enzymes for pro-drugs in the
PT   treatment of cancers
XX
PS   Disclosure; Page 37-39; 64pp; English.
XX
CC   CEA genomic clone lambdaCEA1 was isolated from human chromosome 19
CC   genomic library LL19NL01 (ATCC 57766).  An HindIII fragment of the
CC   clone extended from -14.5 to -10.7 kb relative to the start
CC   site of CEA mRNA.  TRS regions of CEA are used to target e.g.
CC   cytosine deaminase to cancer cells for prodrug activation.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 3774 BP; 906 A; 1066 C; 1057 G; 745 T; 0 other;

  Query Match             19.9%;  Score 94;  DB 16;  Length 3774;
  Best Local Similarity   69.5%;  Pred. No. 5.2e-17;
  Matches  185;  Conservative    0;  Mismatches   50;  Indels   31;  Gaps    3;

Qy          4 CACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTGGGC 63
              ||  |  |||||||    | || ||||||||||||||||||||||||||||||| |||||
Db       1662 CAGTAAACAGTGAGCTCCTATCCAGCCCCCAGAGCCACCTCTGTCACCTTCCTGCTGGGC 1721

Qy         64 ATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGGTTT 123
              |||||||||||||| || ||| ||  ||||||||||||||| ||   |  |||||||||
Db       1722 ATCATCCCACCTTCACA-AGCACTAAAGAGCATGGGGAGACCTGG---CTAGCTGGGTTT 1777

Qy        124 CTCTGTCACAAAGGAAAATAATCCCCCTGGT----------------------------GT 156
              ||   |||||||| |||  |  ||||  | |                            |
Db       1778 CTGCATCACAAAGAAAATAATCCCCCAGGTTCGGATTCCCAGGGCTCTGTATGTGGAGCT 1837

Qy        157 GACAGACCCAAGGACAGAACACAGCAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCC 216
              ||||||||  ||| ||| | | |||||||||||||| || |||| |   |||  ||| |||
Db       1838 GACAGACCTGAGGCCAGGAGATAGCAGAGGTCAGCCCTAGGGAGGGTGGGTCATCCACCC 1897

Qy        217 AGGGGATGGGGGTCCATCCACCTTGC 242
              ||||||  |||||  || |  ||||||
Db       1898 AGGGGACAGGGGTGCACCAGCCTTGC 1923


RESULT 14
ABK83462
ID   ABK83462 standard; cDNA; 39801 BP.
XX
AC   ABK83462;
XX
DT   14-AUG-2002  (first entry)
XX
DE   Human cDNA differentially expressed in granulocytic cells #33.
XX
KW   Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW   viral infection; parasitic infection; protozoal infection;
KW   fungal infection; sterile inflammatory disease; psoriasis;
KW   rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW   cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW   adult respiratory distress syndrome; inflammatory bowel disease;
KW   Crohn's disease; ulcerative colitis; periodontal disease;
```

KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US30821.
XX
PR 03-OCT-2000; 2000US-237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
DR WPI; 2002-435328/46.
XX
PT Detecting granulocyte activation by detecting differential expression
PT of genes associated with granulocyte activation, which serves as
PT diagnostic markers that is useful for monitoring disease states and
PT drug toxicity -
XX
PS Claim 1; SEQ ID No 33; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing
CC the expression level to an expression level in an unactivated
CC GC, where differential expression of Gs is indicative of GCA.
CC Also included are modulating (M2) GA by contacting GC with an agent
CC that alters the expression of at least one gene in Gs; (2) screening (M3)
CC for an agent capable of modulating GCA or an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease using the
CC gene expression profile; (3) detecting (M4) an inflammation (especially
CC chronic) in a tissue, an allergic response in a subject, exposure of a
CC subject to a pathogen or sterile inflammatory disease, by detecting the
CC level of expression in a sample of the tissue of gene(s) from Gs, where
CC the level of expression of the gene is indicative of inflammation;
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,
CC an allergic response in a subject, exposure of a subject to a pathogen
CC or sterile inflammatory disease, by contacting a tissue having
CC inflammation with an agent that modulates the expression of gene(s)
CC from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
CC modulating GA; M3 is useful for screening an agent capable of modulating
CC GCA preferably in an inflammation in a tissue; M4 is useful for
CC detecting an inflammation (especially chronic) in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
CC glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
CC reperfusion injury, ARDS, adult respiratory distress syndrome,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC periodontal disease; also bacterial infection, viral infection,
CC parasitic infection, protozoal infection, fungal infection and M5 is
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 39801 BP; 9969 A; 9367 C; 9531 G; 10934 T; 0 other;

Query Match 18.2%; Score 85.8; DB 24; Length 39801;
Best Local Similarity 61.8%; Pred. No. 3.4e-14;
Matches 201; Conservative 0; Mismatches 97; Indels 27; Gaps 3;

```
Qy       156 TGACAGACCCAAGGACAGAACACAGCAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCC 215
            ||||||| ||||   |   ||||||||||||||    ||||||  ||||||  ||||
Db      9275 TGACAGAGCCAATCCCTAGACACAGCAGAGGTCACTTATGGGGAGCACAGGTCATCCTGT 9334

Qy       216 CAGGGGATGGGGGTCC----ATCCACCTTGCCGAAAAGATTTGTCTGAGGAACTGAAA-- 269
```

```
            | ||||||||||||||    |  |       |  | ||  | |||| ||||||||||
Db      9335 TAAGGGATGGGGGTCCCAATAAGCCTTGCTTCTCACAGCGTGGTCTAAGGAACTGAAATG 9394

Qy       270 ---ATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGAAATGAGAGGGGAGGGGACAG 326
                || | ||  ||| | ||  ||      || ||||||| | |||||||||||| | | |
Db      9395 TAGATTGGAGAAAAAGAGGAAAAGAAGACAAGGAGGCAGGACTGAGAGGGGAGGAGGCGG 9454

Qy       327 AGGACACCTGAATAAAGACCA----------------CACCCATGACCCACGTGATGC 368
             |           |  ||| |                | ||     | | | ||   |
Db      9455 AAAGGCTACCTGTGCAGATCCTCATTGCCACTTACCTTTCCCCGCCCCCACCCCTGGCCC 9514

Qy       369 TGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAGCAGACCA 428
             || ||||| || ||||||| ||| ||| || ||  ||||| |||||||||||||||||| |
Db      9515 TGGGAAGTGCTTCTGCCCTGGGAGGAGGCTTAGCACAGAGAGAGGAAGGACAGCAGAGCT 9574

Qy       429 GACAGTCACAGCAGCCTTGACAAAA 453
             |||||||||||| |||  |||||| |
Db      9575 GACAGTCACAGTAGCTCTGACAAGA 9599
```

RESULT 15
ABN95819
ID ABN95819 standard; DNA; 2974 BP.
XX
AC ABN95819;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2317 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 2317; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 2974 BP; 840 A; 847 C; 613 G; 674 T; 0 other;

Query Match 15.9%; Score 75; DB 24; Length 2974;
Best Local Similarity 100.0%; Pred. No. 1.6e-11;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        397 CTCAGGGCAGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGT 456
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTCAGGGCAGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGT 60

Qy        457 TCCTGGAACTCAAGC 471
              |||||||||||||||
Db         61 TCCTGGAACTCAAGC 75
```

Search completed: September 26, 2003, 12:34:57
Job time : 251 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 26, 2003, 12:22:31 ; Search time 71 Seconds
(without alignments)
2934.265 Million cell updates/sec

Title: US-10-045-116-1
Perfect score: 472
Sequence: 1 AGCCACCACCCAGTGAGCCT.....ACGTTCCTGGAACTCAAGCA 472

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	472	100.0	472	4	US-09-033-556-7	Sequence 7, Appli
	2	471	99.8	11288	3	US-08-646-301A-1	Sequence 1, Appli
	3	471	99.8	11288	4	US-08-481-968A-4	Sequence 4, Appli
	4	471	99.8	11288	4	US-08-154-712B-4	Sequence 4, Appli
	5	471	99.8	15056	4	US-09-474-699-10	Sequence 10, Appl
	6	94	19.9	3774	3	US-08-646-301A-2	Sequence 2, Appli
	7	94	19.9	3774	4	US-08-481-968A-5	Sequence 5, Appli
	8	94	19.9	3774	4	US-08-154-712B-5	Sequence 5, Appli
c	9	71.2	15.1	7218	1	US-08-232-463-14	Sequence 14, Appl
	10	57.6	12.2	387	4	US-09-370-838-177	Sequence 177, App
	11	54	11.4	2220	1	US-08-389-459A-16	Sequence 16, Appl
	12	54	11.4	2220	3	US-08-987-867A-16	Sequence 16, Appl
c	13	44.2	9.4	319	3	US-09-385-982-486	Sequence 486, App
	14	39.2	8.3	289	3	US-09-007-005-17	Sequence 17, Appl
	15	39.2	8.3	289	3	US-09-244-796-17	Sequence 17, Appl
	16	38	8.1	7898	4	US-08-984-709A-49	Sequence 49, Appl
	17	37	7.8	1413	4	US-08-984-709A-52	Sequence 52, Appl
	18	35.6	7.5	50	1	US-08-171-389-392	Sequence 392, App
	19	35.6	7.5	50	1	US-08-123-936-392	Sequence 392, App
	20	35.6	7.5	50	2	US-08-475-228A-392	Sequence 392, App
	21	35.6	7.5	50	3	US-08-482-080A-392	Sequence 392, App
	22	35.6	7.5	50	4	US-09-354-947-392	Sequence 392, App
	23	35.6	7.5	50	5	PCT-US93-12388-392	Sequence 392, App
c	24	34.2	7.2	1338	3	US-09-027-064-3	Sequence 3, Appli
c	25	34.2	7.2	1338	3	US-09-271-815-3	Sequence 3, Appli
c	26	34.2	7.2	2394	3	US-09-027-064-1	Sequence 1, Appli
c	27	34.2	7.2	2394	3	US-09-271-815-1	Sequence 1, Appli
	28	34	7.2	1503	4	US-09-184-418C-29	Sequence 29, Appl
	29	34	7.2	8953	4	US-09-184-418C-3	Sequence 3, Appli
	30	33.8	7.2	2397	1	US-07-891-942G-11	Sequence 11, Appl
	31	33.4	7.1	42325	4	US-08-311-731A-131	Sequence 131, App
	32	33.2	7.0	248	3	US-09-007-005-32	Sequence 32, Appl
	33	33.2	7.0	248	3	US-09-244-796-32	Sequence 32, Appl
	34	33.2	7.0	277	3	US-09-007-005-3	Sequence 3, Appli
	35	33.2	7.0	277	3	US-09-244-796-3	Sequence 3, Appli
	36	32.6	6.9	3995	3	US-08-868-435-3	Sequence 3, Appli
	37	32.6	6.9	3995	4	US-08-744-231-3	Sequence 3, Appli
	38	32.2	6.8	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	39	32	6.8	1298	3	US-08-948-705-3	Sequence 3, Appli
	40	32	6.8	1298	4	US-09-510-543-3	Sequence 3, Appli
	41	32	6.8	6755	3	US-08-931-999-4	Sequence 4, Appli
	42	31.8	6.7	3383	1	US-07-707-367-1	Sequence 1, Appli
	43	31.8	6.7	3989	4	US-09-205-258-28	Sequence 28, Appl
	44	31.8	6.7	4117	4	US-09-484-970B-2	Sequence 2, Appli
c	45	31.6	6.7	486	4	US-09-510-252-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-09-033-556-7
; Sequence 7, Application US/09033556
; Patent No. 6432700
;  GENERAL INFORMATION:
;    APPLICANT:  Henderson, Daniel R.
;    APPLICANT:  Yu, De Chao
;    TITLE OF INVENTION:  ADENOVIRUS VECTORS CONTAINING
;    TITLE OF INVENTION:  HETEROLOGOUS TRANSCRIPTION REGULATORY ELEMENTS AND METHODS
;    TITLE OF INVENTION:  OF USING SAME
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/033,556
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Catherine, Polizzi M
;      REGISTRATION NUMBER:  40,130
;      REFERENCE/DOCKET NUMBER:  34802-20010.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-813-5600
;      TELEFAX:  650-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  472 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-033-556-7

  Query Match           100.0%;  Score 472;  DB 4;  Length 472;
  Best Local Similarity  100.0%;  Pred. No. 6.3e-129;
  Matches  472;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
```

RESULT 2
US-08-646-301A-1
; Sequence 1, Application US/08646301A
; Patent No. 6194211
; GENERAL INFORMATION:
; APPLICANT: Richards, Cynthia Ann
; APPLICANT: Huber, Brian E.
; TITLE OF INVENTION: Transcriptional Regulatory Sequence of Carcinoembryonic
; Patent No. 6194211
; TITLE OF INVENTION: Antigen for Expression Targeting
; FILE REFERENCE: PB1508USW
; CURRENT APPLICATION NUMBER: US/08/646,301A
; CURRENT FILING DATE: 1996-05-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11288
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-646-301A-1

Query Match 99.8%; Score 471; DB 3; Length 11288;
Best Local Similarity 100.0%; Pred. No. 4.4e-128;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10294 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 10353

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10354 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 10413

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10414 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 10473

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10474 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 10533

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10534 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 10593

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10594 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 10653

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10654 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 10713

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10714 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 10764
```

RESULT 3
US-08-481-968A-4
; Sequence 4, Application US/08481968A
; Patent No. 6300490
; GENERAL INFORMATION:
; APPLICANT: Huber, Brian
; APPLICANT: Richards, Cynthia
; TITLE OF INVENTION: Molecular Constructs Comprising a Carcinoembryonic Antigen
; TITLE OF INVENTION: Transcriptional Regulatory Region
; FILE REFERENCE: PB1087US4
; CURRENT APPLICATION NUMBER: US/08/481,968A
; CURRENT FILING DATE: 1998-06-07
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 11288
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-481-968A-4

Query Match 99.8%; Score 471; DB 4; Length 11288;
Best Local Similarity 100.0%; Pred. No. 4.4e-128;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10294 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 10353

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10354 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 10413

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10414 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 10473

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10474 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 10533

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10534 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 10593

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10594 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 10653

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10654 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 10713

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10714 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 10764
```

```
RESULT 4
US-08-154-712B-4
; Sequence 4, Application US/08154712B
; Patent No. 6337209
; GENERAL INFORMATION:
;  APPLICANT:  Huber, Brian
;  APPLICANT:  Richards, Cynthia
;  TITLE OF INVENTION:  Molecular Constructs Containing a Carcinoembryonic Antigen Regu
;  TITLE OF INVENTION:  Sequence
;  FILE REFERENCE:  PB1087US3
;  CURRENT APPLICATION NUMBER: US/08/154,712B
;  CURRENT FILING DATE:  1993-11-19
;  NUMBER OF SEQ ID NOS:  36
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 4
;  LENGTH: 11288
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-08-154-712B-4

  Query Match             99.8%;  Score 471;  DB 4;  Length 11288;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-128;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10294 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 10353

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10354 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 10413

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10414 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 10473

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10474 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 10533

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10534 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 10593

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10594 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 10653

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10654 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 10713

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10714 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 10764


RESULT 5
US-09-474-699-10
; Sequence 10, Application US/09474699
; Patent No. 6495130
; GENERAL INFORMATION:
;  APPLICANT:  Henderson, Daniel R.
;  APPLICANT:  Yu, De Chao
;  TITLE OF INVENTION:  TARGET CELL-SPECIFIC ADENOVIRAL VECTORS
;  TITLE OF INVENTION:  CONTAINING E3 AND METHODS OF USE THEREOF
;  FILE REFERENCE:  348022001300
;  CURRENT APPLICATION NUMBER: US/09/474,699
;  CURRENT FILING DATE:  1999-12-29
;  PRIOR APPLICATION NUMBER:  60/114,262
;  PRIOR FILING DATE:  1998-12-30
;  NUMBER OF SEQ ID NOS:  23
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;  LENGTH: 15056
;  TYPE: DNA
;  ORGANISM: Homo Sapien
US-09-474-699-10

  Query Match             99.8%;  Score 471;  DB 4;  Length 15056;
  Best Local Similarity   100.0%;  Pred. No. 5e-128;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532


RESULT 6
US-08-646-301A-2
; Sequence 2, Application US/08646301A
; Patent No. 6194211
; GENERAL INFORMATION:
;  APPLICANT:  Richards, Cynthia Ann
;  APPLICANT:  Huber, Brian E.
;  TITLE OF INVENTION:  Iranscriptional Regulatory Sequence of Carcinoembryonic
; Patent No. 6194211
;  TITLE OF INVENTION:  Antigen for Expression Targeting
;  FILE REFERENCE:  PB1508USW
;  CURRENT APPLICATION NUMBER: US/08/646,301A
;  CURRENT FILING DATE:  1996-05-16
;  NUMBER OF SEQ ID NOS:  25
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 2
;  LENGTH: 3774
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-08-646-301A-2

  Query Match             19.9%;  Score 94;  DB 3;  Length 3774;
  Best Local Similarity   69.5%;  Pred. No. 6.7e-18;
  Matches  185;  Conservative    0;  Mismatches   50;  Indels   31;  Gaps    3;

Qy          4 CACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCICTGTCACCTTCCTGTTGGGC 63
              ||  |  ||||||||   | || ||||||||||||||||||||||||||||| |||||
Db       1662 CAGTAAACAGTGAGCTCCTATCCAGCCCCCAGAGCCACCTCTGTCACCTTCCTGCTGGGC 1721
```

```
Qy          64 ATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGGTTT 123
               ||||||||||||||| || ||| ||  |||||||||||||||| ||   |  |||||||||
Db        1722 ATCATCCCACCTTCACA-AGCACTAAAGAGCATGGGGAGACCTGG---CTAGCTGGGTTT 1777

Qy         124 CTCTGTCACAAAGGAAAATAATCCCCCTGGT--------------------------GT 156
               ||   |||||||| |||  |  ||||  | |                          |
Db        1778 CTGCATCACAAAGAAAATAATCCCCCAGGTTCGGATTCCCAGGGCTCTGTATGTGGAGCT 1837

Qy         157 GACAGACCCAAGGACAGAACACAGCAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCC 216
               ||||||||  ||| ||| | | ||||||||||||| || |||| |   |||  ||| |||
Db        1838 GACAGACCTGAGGCCAGGAGATAGCAGAGGTCAGCCCTAGGGAGGGTGGGTCATCCACCC 1897

Qy         217 AGGGGATGGGGGTCCATCCACCTTGC 242
               ||||||  ||||| || |  ||||||
Db        1898 AGGGGACAGGGGTGCACCAGCCTTGC 1923


RESULT 7
US-08-481-968A-5
; Sequence 5, Application US/08481968A
; Patent No. 6300490
; GENERAL INFORMATION:
;  APPLICANT:  Huber, Brian
;  APPLICANT:  Richards, Cynthia
;  TITLE OF INVENTION:  Molecular Constructs Comprising a Carcinoembryonic Antigen (CEA
;  TITLE OF INVENTION:  Transcriptional Regulatory Region
;  FILE REFERENCE:  PB1087US4
;  CURRENT APPLICATION NUMBER: US/08/481,968A
;  CURRENT FILING DATE:  1998-06-07
;  NUMBER OF SEQ ID NOS:  36
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 5
;  LENGTH: 3774
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-08-481-968A-5

  Query Match             19.9%;  Score 94;  DB 4;  Length 3774;
  Best Local Similarity   69.5%;  Pred. No. 6.7e-18;
  Matches  185;  Conservative    0;  Mismatches   50;  Indels   31;  Gaps    3;

Qy           4 CACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTGGGC 63
               ||  |  ||||||||   | || ||||||||||||||||||||||||||||||| |||||
Db        1662 CAGTAAACAGTGAGCTCCTATCCAGCCCCCAGAGCCACCTCTGTCACCTTCCTGCTGGGC 1721

Qy          64 ATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGGTTT 123
               ||||||||||||||| || ||| ||  |||||||||||||||| ||   |  |||||||||
Db        1722 ATCATCCCACCTTCACA-AGCACTAAAGAGCATGGGGAGACCTGG---CTAGCTGGGTTT 1777

Qy         124 CTCTGTCACAAAGGAAAATAATCCCCCTGGT--------------------------GT 156
               ||   |||||||| |||  |  ||||  | |                          |
Db        1778 CTGCATCACAAAGAAAATAATCCCCCAGGTTCGGATTCCCAGGGCTCTGTATGTGGAGCT 1837

Qy         157 GACAGACCCAAGGACAGAACACAGCAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCC 216
               ||||||||  ||| ||| | | ||||||||||||| || |||| |   |||  ||| |||
Db        1838 GACAGACCTGAGGCCAGGAGATAGCAGAGGTCAGCCCTAGGGAGGGTGGGTCATCCACCC 1897

Qy         217 AGGGGATGGGGGTCCATCCACCTTGC 242
               ||||||  ||||| || |  ||||||
Db        1898 AGGGGACAGGGGTGCACCAGCCTTGC 1923


RESULT 8
US-08-154-712B-5
; Sequence 5, Application US/08154712B
; Patent No. 6337209
; GENERAL INFORMATION:
;  APPLICANT:  Huber, Brian
;  APPLICANT:  Richards, Cynthia
;  TITLE OF INVENTION:  Molecular Constructs Containing a Carcinoembryonic Antigen Regu
;  TITLE OF INVENTION:  Sequence
;  FILE REFERENCE:  PB1087US3
;  CURRENT APPLICATION NUMBER: US/08/154,712B
;  CURRENT FILING DATE:  1993-11-19
;  NUMBER OF SEQ ID NOS:  36
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 5
;  LENGTH: 3774
;  TYPE: DNA
;  ORGANISM: Homo sapiens
US-08-154-712B-5

  Query Match             19.9%;  Score 94;  DB 4;  Length 3774;
  Best Local Similarity   69.5%;  Pred. No. 6.7e-18;
  Matches  185;  Conservative    0;  Mismatches   50;  Indels   31;  Gaps    3;

Qy           4 CACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTGGGC 63
               ||  |  ||||||||   | || ||||||||||||||||||||||||||||||| |||||
Db        1662 CAGTAAACAGTGAGCTCCTATCCAGCCCCCAGAGCCACCTCTGTCACCTTCCTGCTGGGC 1721

Qy          64 ATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGGTTT 123
               ||||||||||||||| || ||| ||  |||||||||||||||| ||   |  |||||||||
Db        1722 ATCATCCCACCTTCACA-AGCACTAAAGAGCATGGGGAGACCTGG---CTAGCTGGGTTT 1777

Qy         124 CTCTGTCACAAAGGAAAATAATCCCCCTGGT--------------------------GT 156
               ||   |||||||| |||  |  ||||  | |                          |
Db        1778 CTGCATCACAAAGAAAATAATCCCCCAGGTTCGGATTCCCAGGGCTCTGTATGTGGAGCT 1837

Qy         157 GACAGACCCAAGGACAGAACACAGCAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCC 216
               ||||||||  ||| ||| | | ||||||||||||| || |||| |   |||  ||| |||
Db        1838 GACAGACCTGAGGCCAGGAGATAGCAGAGGTCAGCCCTAGGGAGGGTGGGTCATCCACCC 1897

Qy         217 AGGGGATGGGGGTCCATCCACCTTGC 242
               ||||||  ||||| || |  ||||||
Db        1898 AGGGGACAGGGGTGCACCAGCCTTGC 1923


RESULT 9
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
```

```
;       TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  7218 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;       CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             15.1%;  Score 71.2;  DB 1;  Length 7218;
  Best Local Similarity   4.0%;  Pred. No. 4.1e-11;
  Matches   16;  Conservative  236;  Mismatches  144;  Indels    0;  Gaps    0;

Qy         76 TCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGGTTTCTCTGTCACAAA 135
              | | ::::    :::::: : :::::::   ::::    :  :::          :  : :::
Db       1437 TACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1378

Qy        136 GGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAGCAGAGGTCAGCACTG 195
              :::::: ::       :: : :: :::   ::::: :::: : :: ::::::  :: :  :
Db       1377 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1318

Qy        196 GGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTTGCCGAAAAGATTTGT 255
              :::::::: :::  :        :::::: :::::   :   :    :  :::::::   :
Db       1317 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1258

Qy        256 CTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGAAATGAGAGG 315
                ::::::  ::::: :::::::::::::::::::: :::::::: ::::: ::::::
Db       1257 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1198

Qy        316 GGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCACGTGATGCTGAGAAG 375
              :::::::: :::::: :   ::: :::::  : :    : ::   : : :: :  ::::::
Db       1197 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1138

Qy        376 TACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAGCAGACCAGACAGTC 435
              :    :    :::::::::   ::::::::::::::: :: :::  ::: ::
Db       1137 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1078

Qy        436 ACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              : :: ::      |  ||  | ||| || ||  |||
Db       1077 RRRRRRRRRRRATCGCAAGCTCCCTCGACCTGCAGC 1042


RESULT 10
US-09-370-838-177
; Sequence 177, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
;  APPLICANT: Reed, Steven G.
;  APPLICANT: Lodes, Michael J.
;  APPLICANT: Mohamath, Roadoh
;  APPLICANT: Secrist, Heather
;  TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
;  TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
;  FILE REFERENCE: 210121.475C1
;  CURRENT APPLICATION NUMBER: US/09/370,838
;  CURRENT FILING DATE: 1999-08-09
;  EARLIER APPLICATION NUMBER: US 09/285,323
;  EARLIER FILING DATE: 1999-04-02
;  NUMBER OF SEQ ID NOS: 289
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 177
;   LENGTH: 387
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-370-838-177

  Query Match             12.2%;  Score 57.6;  DB 4;  Length 387;
  Best Local Similarity   93.8%;  Pred. No. 1.2e-07;
  Matches   60;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        408 GGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTC 467
              || |  | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 GGCACCAGGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTC 66

Qy        468 AAGC 471
              ||||
Db         67 AAGC 70


RESULT 11
US-08-389-459A-16
; Sequence 16, Application US/08389459A
; Patent No. 5817512
;  GENERAL INFORMATION:
;    APPLICANT:  Morrow, Casey D. and Porter, Donna, C.
;    TITLE OF INVENTION:  ENCAPSIDATED RECOMBINANT POLIOVIRUS
;    TITLE OF INVENTION:  NUCLEIC ACID AND METHODS OF MAKING AND
;    TITLE OF INVENTION:  USING SAME
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 STATE SIREET, SUITE 510
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/389,459A
;      FILING DATE:  15-FEB-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/087,009
;      FILING DATE:  01-JUL-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Silveri, Jean M.
;      REGISTRATION NUMBER:  39,030
;      REFERENCE/DOCKET NUMBER:  UAG-004CP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2220 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..2203
US-08-389-459A-16

  Query Match             11.4%;  Score 54;  DB 1;  Length 2220;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-06;
  Matches   54;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        419 CAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 CAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 58


RESULT 12
US-08-987-867A-16
; Sequence 16, Application US/08987867A
; Patent No. 6063384
;  GENERAL INFORMATION:
;    APPLICANT:  C. Morrow et al.
;    TITLE OF INVENTION:  ENCAPSIDAIED RECOMBINANT VIRAL
```

```
;    TITLE OF INVENTION:  NUCLEIC ACID AND METHODS OF MAKING AND
;    TITLE OF INVENTION:  USING SAME
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  28 STATE STREET
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/987,867A
;      FILING DATE:  09-DEC-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/087,009
;      FILING DATE:  01-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Myers, Louis
;      REGISTRATION NUMBER:  35,965
;      REFERENCE/DOCKET NUMBER:  UAP-004CPDV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 742-4214
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2220 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..2203
US-08-987-867A-16

  Query Match             11.4%;  Score 54;  DB 3;  Length 2220;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-06;
  Matches   54;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        419 CAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 CAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 58


RESULT 13
US-09-385-982-486/c
; Sequence 486, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
;  APPLICANT: ENDEGE, WILSON O., ET AL.
;  TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
;  TITLE OF INVENTION: PRODUCTS:  II
;  FILE REFERENCE: CCDNA-260XX
;  CURRENT APPLICATION NUMBER: US/09/385,982
;  CURRENT FILING DATE: 1999-08-30
;  EARLIER APPLICATION NUMBER: 09/328,111
;  EARLIER FILING DATE: 1999-06-08
;  EARLIER APPLICATION NUMBER: 60/117,393
;  EARLIER FILING DATE: 1999-01-27
;  EARLIER APPLICATION NUMBER: 60/098,639
;  EARLIER FILING DATE: 1998-08-31
;  NUMBER OF SEQ ID NOS: 544
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 486
;   LENGTH: 319
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
US-09-385-982-486

  Query Match             9.4%;  Score 44.2;  DB 3;  Length 319;
  Best Local Similarity   87.0%;  Pred. No. 0.00094;
  Matches   60;  Conservative    0;  Mismatches    8;  Indels    1;  Gaps    1;

Qy        404 CAGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAAC-GTTCCTGG 462
              |||||| ||||||||||||||| | ||||||||||||||| |||||| | ||||||||
Db        309 CAGAGGTAGGAAGGACAGCAGAGCTGACAGTCACAGCAGCCCTGACAAGAGAGTTCCTGG 250

Qy        463 AACTCAAGC 471
              | | |||||
Db        249 AGCCCAAGC 241


RESULT 14
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350003
;  CURRENT APPLICATION NUMBER: US/09/007,005B
;  CURRENT FILING DATE: 1998-01-14
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

  Query Match             8.3%;  Score 39.2;  DB 3;  Length 289;
  Best Local Similarity   3.9%;  Pred. No. 0.026;
  Matches    8;  Conservative   97;  Mismatches  101;  Indels    0;  Gaps    0;

Qy        140 AATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAGCAGAGGTCAGCACTGGGGA 199
              :| |:      ::    ::: :  :: : ::: :     : ::: :   : :  :: :
Db         49 RARARURGRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 108

Qy        200 AGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTTGCCGAAAAGATTTGTCTGA 259
              ::: : :           ::: : ::: :  ::   ::     :: : :::   :  :
Db        109 RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 168

Qy        260 GGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGAAATGAGAGGGGAG 319
              :::   : : : ::: : : : : ::: : :::   : : : ::: : : : :|:|: :
Db        169 RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRCRARGRCRU 228

Qy        320 GGGACAGAGGACACCTGAATAAAGAC 345
              :|:   : : : :| :: ::: :|:|
Db        229 RGRCRGRURARARCRURCRURURGRC 254


RESULT 15
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
```

```
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350007
;  CURRENT APPLICATION NUMBER: US/09/244,796
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  EARLIER APPLICATION NUMBER: 09/007,005
;  EARLIER FILING DATE: 1998-01-14
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

  Query Match             8.3%;  Score 39.2;  DB 3;  Length 289;
  Best Local Similarity   3.9%;  Pred. No. 0.026;
  Matches    8;  Conservative   97;  Mismatches  101;  Indels    0;  Gaps    0;

Qy        140 AATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAGCAGAGGTCAGCACTGGGGA 199
              :| |::      ::    ::: :  :: : ::: :     : ::: :   : :  :: :
Db         49 RARARURGRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 108

Qy        200 AGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTTGCCGAAAAGATTTGTCTGA 259
              ::: : :           ::: : ::: :  ::   ::     :: : :::   :   :
Db        109 RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRN 168

Qy        260 GGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGAAATGAGAGGGGAG 319
              :::   : : : ::: : : : : ::: : :::   : : : ::: : : : :|:|: :
Db        169 RSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRNRNRSRCRARGRCRU 228

Qy        320 GGGACAGAGGACACCTGAATAAAGAC 345
              :|:   : : : :| :: ::: :|:|
Db        229 RGRCRGRURARARCRURCRURURGRC 254


Search completed: September 26, 2003, 13:42:36
Job time : 74 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 26, 2003, 13:07:01 ; Search time 207 Seconds
(without alignments)
5676.520 Million cell updates/sec

Title: US-10-045-116-1
Perfect score: 472
Sequence: 1 AGCCACCACCCAGTGAGCCT.....ACGTTCCTGGAACTCAAGCA 472

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1678620 seqs, 1244745471 residues

Total number of hits satisfying chosen parameters: 3357240

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	472	100.0	472	11	US-09-151-376-7	Sequence 7, Appli
2	472	100.0	472	11	US-09-151-376-54	Sequence 54, Appl
3	472	100.0	472	12	US-10-139-089-7	Sequence 7, Appli
4	472	100.0	472	12	US-10-139-089-54	Sequence 54, Appl
5	472	100.0	472	14	US-10-045-116-1	Sequence 1, Appli
6	471	99.8	11288	9	US-09-947-925A-4	Sequence 4, Appli
7	471	99.8	11288	10	US-09-822-634-1	Sequence 1, Appli
8	471	99.8	15056	9	US-09-392-822-4	Sequence 4, Appli
9	471	99.8	15056	11	US-09-814-357-14	Sequence 14, Appl
10	471	99.8	15056	12	US-09-814-351-14	Sequence 14, Appl
11	471	99.8	15056	14	US-10-045-116-25	Sequence 25, Appl
12	471	99.8	15056	14	US-10-226-820-10	Sequence 10, Appl
13	432.4	91.6	3281	12	US-10-090-238-1	Sequence 1, Appli
14	350.4	74.2	425	14	US-10-212-667-2	Sequence 2, Appli
15	120.8	25.6	1855	13	US-10-027-632-97812	Sequence 97812, A
16	99.2	21.0	608	13	US-10-027-632-8715	Sequence 8715, Ap
17	94	19.9	3774	9	US-09-947-925A-5	Sequence 5, Appli
18	75	15.9	2974	10	US-09-954-456-56	Sequence 56, Appl
19	75	15.9	2974	10	US-09-880-107-2317	Sequence 2317, Ap
20	75	15.9	2974	14	US-10-157-031-340	Sequence 340, App
21	75	15.9	2974	14	US-10-207-655-86	Sequence 86, Appl
22	67.8	14.4	2547	12	US-09-814-353-20321	Sequence 20321, A
23	60.6	12.8	474	9	US-09-864-761-6512	Sequence 6512, Ap
24	59	12.5	1298	9	US-09-925-301-290	Sequence 290, App
25	57.6	12.2	387	10	US-09-738-973-177	Sequence 177, App
26	57.6	12.2	387	10	US-09-854-133-177	Sequence 177, App
27	57.6	12.2	387	14	US-10-144-649A-177	Sequence 177, App
28	57.2	12.1	460	9	US-09-922-217-169	Sequence 169, App
29	57.2	12.1	460	10	US-09-833-263-169	Sequence 169, App
30	57.2	12.1	460	13	US-10-025-380-169	Sequence 169, App
c 31	56	11.9	676	13	US-10-027-632-127633	Sequence 127633,
c 32	55.4	11.7	294	10	US-09-920-300A-311	Sequence 311, App
c 33	55.4	11.7	294	12	US-10-099-926-311	Sequence 311, App
c 34	55.4	11.7	294	13	US-10-033-528-311	Sequence 311, App
35	54	11.4	2220	9	US-09-756-551A-16	Sequence 16, Appl
c 36	52	11.0	288	9	US-09-815-343-1478	Sequence 1478, Ap
37	51.6	10.9	1833	12	US-09-814-353-20347	Sequence 20347, A
38	51.6	10.9	1833	14	US-10-198-846-13797	Sequence 13797, A
c 39	49.2	10.4	5975	12	US-10-311-455-210	Sequence 210, App
40	47	10.0	283	9	US-09-815-343-349	Sequence 349, App
c 41	44.2	9.4	319	11	US-09-871-161-486	Sequence 486, App
c 42	42.4	9.0	409	12	US-09-814-353-16466	Sequence 16466, A
43	42.4	9.0	627	14	US-10-198-846-7704	Sequence 7704, Ap
44	42.4	9.0	2134	14	US-10-208-408-5	Sequence 5, Appli
45	42.4	9.0	2364	14	US-10-198-846-10292	Sequence 10292, A

ALIGNMENTS

RESULT 1
US-09-151-376-7
; Sequence 7, Application US/09151376
; Publication No. US20030044383A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, D.R.
; APPLICANT: Schuur, E.R.
; TITLE OF INVENTION: TISSUE SPECIFIC VIRAL VECTORS
; FILE REFERENCE: 348022000221
; CURRENT APPLICATION NUMBER: US/09/151,376
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: 08/669,753
; EARLIER FILING DATE: 1996-06-26
; EARLIER APPLICATION NUMBER: 08/495,034
; EARLIER FILING DATE: 1995-06-27
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-151-376-7

Query Match 100.0%; Score 472; DB 11; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.7e-137;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
```

```
Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
```

RESULT 2
US-09-151-376-54
; Sequence 54, Application US/09151376
; Publication No. US20030044383A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, D.R.
; APPLICANT: Schuur, E.R.
; TITLE OF INVENTION: TISSUE SPECIFIC VIRAL VECTORS
; FILE REFERENCE: 348022000221
; CURRENT APPLICATION NUMBER: US/09/151,376
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: 08/669,753
; EARLIER FILING DATE: 1996-06-26
; EARLIER APPLICATION NUMBER: 08/495,034
; EARLIER FILING DATE: 1995-06-27
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Unknown
US-09-151-376-54

Query Match 100.0%; Score 472; DB 11; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.7e-137;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCACCACCCAGTGAGCCTTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
```

RESULT 3
US-10-139-089-7
; Sequence 7, Application US/10139089
; Publication No. US20030152553A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, D.R.
; APPLICANT: Schuur, E.R.
; TITLE OF INVENTION: TISSUE SPECIFIC VIRAL VECTORS
; FILE REFERENCE: 348022000221
; CURRENT APPLICATION NUMBER: US/10/139,089
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 08/669,753
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: 08/495,034
; PRIOR FILING DATE: 1995-06-27
; PRIOR APPLICATION NUMBER: 09/509,591
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 09/151,376
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 09/033,428
; PRIOR FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: 60/039,597
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: 09/033,555
; PRIOR FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: 60/039,763
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: 09/033,333
; PRIOR FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: 60/039,762
; PRIOR FILING DATE: 1997-03-03
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-139-089-7

Query Match 100.0%; Score 472; DB 12; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.7e-137;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCACCACCCAGTGAGCCTTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
```

```
Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
```

RESULT 4
US-10-139-089-54
; Sequence 54, Application US/10139089
; Publication No. US20030152553A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, D.R.
; APPLICANT: Schuur, E.R.
; TITLE OF INVENTION: TISSUE SPECIFIC VIRAL VECTORS
; FILE REFERENCE: 348022000221
; CURRENT APPLICATION NUMBER: US/10/139,089
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 08/669,753
; PRIOR FILING DATE: 1996-06-26
; PRIOR APPLICATION NUMBER: 08/495,034
; PRIOR FILING DATE: 1995-06-27
; PRIOR APPLICATION NUMBER: 09/509,591
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 09/151,376
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 09/033,428
; PRIOR FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: 60/039,597
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: 09/033,555
; PRIOR FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: 60/039,763
; PRIOR FILING DATE: 1997-03-03
; PRIOR APPLICATION NUMBER: 09/033,333
; PRIOR FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: 60/039,762
; PRIOR FILING DATE: 1997-03-03
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 472
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Unknown
US-10-139-089-54

Query Match 100.0%; Score 472; DB 12; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.7e-137;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCIT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
```

RESULT 5
US-10-045-116-1
; Sequence 1, Application US/10045116
; Publication No. US20030026792A1
; GENERAL INFORMATION:
; APPLICANT: Lamparski, Henry
; Schuur, Eric
; Henderson, Daniel
; TITLE OF INVENTION: ADENOVIRUS VECTORS SPECIFIC FOR CELLS
; EXPRESSING CARCINOEMBRYONIC ANTIGENS AND METHODS OF USE
; THEREOF
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/045,116
; FILING DATE: 23-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/[illegible],555A
; FILING DATE: 02-Mar-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: POLIZZI, CATHERINE M.
; REGISTRATION NUMBER: 40,130
; REFERENCE/DOCKET NUMBER: 34802-30005.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141 MRSNFOERS SFO
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-045-116-1

Query Match 100.0%; Score 472; DB 14; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.7e-137;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
```

```
Qy        61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120

Qy       121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180

Qy       181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240

Qy       241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300

Qy       301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360

Qy       361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420

Qy       421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGCA 472
```

RESULT 6
US-09-947-925A-4
; Sequence 4, Application US/09947925A
; Patent No. US20020055482A1
; GENERAL INFORMATION:
; APPLICANT: Huber, Brian
; APPLICANT: Richards, Cynthia
; TITLE OF INVENTION: Molecular Constructs Containing a Carcinoembryonic
; TITLE OF INVENTION: Antigen Regulatory
; TITLE OF INVENTION: Sequence
; FILE REFERENCE: PB1087US3
; CURRENT APPLICATION NUMBER: US/09/947,925A
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US/08/154,712
; PRIOR FILING DATE: 1993-11-19
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 11288
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-947-925A-4

Query Match 99.8%; Score 471; DB 9; Length 11288;
Best Local Similarity 100.0%; Pred. No. 2.7e-136;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10294 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 10353

Qy        61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10354 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 10413

Qy       121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10414 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 10473

Qy       181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10474 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 10533

Qy       241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10534 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 10593

Qy       301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10594 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 10653

Qy       361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10654 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 10713

Qy       421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10714 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 10764
```

RESULT 7
US-09-822-634-1
; Sequence 1, Application US/09822634
; Patent No. US20020150556A1
; GENERAL INFORMATION:
; APPLICANT: Vile, Richard G.
; APPLICANT: Harrington, Kevin
; APPLICANT: Bateman, Andrew
; APPLICANT: Murphy, Steven
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TISSUE
; TITLE OF INVENTION: SPECIFIC GENE REGULATION THERAPY
; FILE REFERENCE: 07039-289001
; CURRENT APPLICATION NUMBER: US/09/822,634
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/193,977
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 11288
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10804)...(10935)
US-09-822-634-1

Query Match 99.8%; Score 471; DB 10; Length 11288;
Best Local Similarity 100.0%; Pred. No. 2.7e-136;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10294 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 10353

Qy        61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10354 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 10413

Qy       121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10414 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 10473

Qy       181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10474 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 10533

Qy       241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10534 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 10593

Qy       301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     10594 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 10653

Qy       361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
```

```
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10654 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 10713

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
                  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10714 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 10764


RESULT 8
US-09-392-822-4
; Sequence 4, Application US/09392822
; Patent No. US20010053352A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, De Chao
;  APPLICANT: Henderson, Daniel
;  TITLE OF INVENTION: Adenovirus Vectors Containing Cell
;  TITLE OF INVENTION: Status-Specific Response Elements and Methods of Use Thereof
;  FILE REFERENCE: 348022001200
;  CURRENT APPLICATION NUMBER: US/09/392,822
;  CURRENT FILING DATE: 1999-09-09
;  EARLIER APPLICATION NUMBER: 60/099,791
;  EARLIER FILING DATE: 1998-09-10
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 15056
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-09-392-822-4

  Query Match             99.8%;  Score 471;  DB 9;  Length 15056;
  Best Local Similarity   100.0%;  Pred. No. 3e-136;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532


RESULT 9
US-09-814-357-14
; Sequence 14, Application US/09814357
; Publication No. US20030068307A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, De-Chao
;  APPLICANT:  Chen, Yu
;  APPLICANT:  Henderson, Daniel R.
;  TITLE OF INVENTION: METHODS OF TREATING NEOPLASIA
;  TITLE OF INVENTION:  WITH COMBINATION TARGET CELL-SPECIFIC ADENOVIRUS,
;  TITLE OF INVENTION:  CHEMOTHERAPY AND RADIATION
;  FILE REFERENCE: 348022001600
;  CURRENT APPLICATION NUMBER: US/09/814,357
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: 60/192,015
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 15056
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: CEA TRE
US-09-814-357-14

  Query Match             99.8%;  Score 471;  DB 11;  Length 15056;
  Best Local Similarity   100.0%;  Pred. No. 3e-136;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532


RESULT 10
US-09-814-351-14
; Sequence 14, Application US/09814351
; Publication No. US20030148520A1
; GENERAL INFORMATION:
;  APPLICANT: Yu, De-Chao
;  APPLICANT:  Li, Yuanhao
;  APPLICANT:  Henderson, Daniel R.
;  TITLE OF INVENTION: CELL-SPECIFIC ADENOVIRUS VECTORS
;  TITLE OF INVENTION:  COMPRISING AN INTERNAL RIBOSOME ENTRY SITE
;  FILE REFERENCE: 348022001700
;  CURRENT APPLICATION NUMBER: US/09/814,351
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/192,156
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 35
```

```
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 15056
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: CEA TRE
US-09-814-351-14

  Query Match             99.8%;  Score 471;  DB 12;  Length 15056;
  Best Local Similarity   100.0%;  Pred. No. 3e-136;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532


RESULT 11
US-10-045-116-25
; Sequence 25, Application US/10045116
; Publication No. US20030026792A1
;   GENERAL INFORMATION:
;        APPLICANT: Lamparski, Henry
;                   Schuur, Eric
;                   Henderson, Daniel
;        TITLE OF INVENTION: ADENOVIRUS VECTORS SPECIFIC FOR CELLS
;                            EXPRESSING CARCINOEMBRYONIC ANTIGENS AND METHODS OF USE
;                            THEREOF
;        NUMBER OF SEQUENCES: 25
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MORRISON & FOERSTER
;             STREET: 755 PAGE MILL ROAD
;             CITY: PALO ALTO
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304-1018
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/045,116
;             FILING DATE: 23-Oct-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/033,555A
;             FILING DATE: 02-Mar-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: POLIZZI, CATHERINE M.
;             REGISTRATION NUMBER: 40,130
;             REFERENCE/DOCKET NUMBER: 34802-30005.00
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (415) 813-5600
;             TELEFAX: (415) 494-0792
;             TELEX: 706141 MRSNFOERS SFO
;   INFORMATION FOR SEQ ID NO: 25:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 15056 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-10-045-116-25

  Query Match             99.8%;  Score 471;  DB 14;  Length 15056;
  Best Local Similarity   100.0%;  Pred. No. 3e-136;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532


RESULT 12
US-10-226-820-10
; Sequence 10, Application US/10226820
; Publication No. US20030118555A1
; GENERAL INFORMATION:
;  APPLICANT: Henderson, Daniel R.
;  APPLICANT:  Yu, De Chao
;  TITLE OF INVENTION: TARGET CELL-SPECIFIC ADENOVIRAL VECTORS
;  TITLE OF INVENTION:  CONTAINING E3 AND METHODS OF USE THEREOF
;  FILE REFERENCE: 348022001300
;  CURRENT APPLICATION NUMBER: US/10/226,820
;  CURRENT FILING DATE:  2002-08-21
;  PRIOR APPLICATION NUMBER: US/09/474,699
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: 60/114,262
```

```
;  PRIOR FILING DATE: 1998-12-30
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 15056
;   TYPE: DNA
;   ORGANISM: Homo Sapien
US-10-226-820-10

  Query Match             99.8%;  Score 471;  DB 14;  Length 15056;
  Best Local Similarity   100.0%;  Pred. No. 3e-136;
  Matches  471;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14062 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 14121

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14122 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 14181

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14182 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 14241

Qy        181 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14242 CAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTT 14301

Qy        241 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14302 GCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGA 14361

Qy        301 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14362 GGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCCA 14421

Qy        361 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14422 CGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACA 14481

Qy        421 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      14482 GCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 14532


RESULT 13
US-10-090-238-1
; Sequence 1, Application US/10090238
; Publication No. US20030176377A1
; GENERAL INFORMATION:
;  APPLICANT: The Scripps Research Institute
;  APPLICANT:  Rong Xiang
;  APPLICANT:  Ralph A. Reisfeld
;  TITLE OF INVENTION: DNA VACCINES ENCODING CEA AND A CD40
;  TITLE OF INVENTION:  LIGAND AND METHODS OF USE THEREOF
;  FILE REFERENCE: TSRI-830.0
;  CURRENT APPLICATION NUMBER: US/10/090,238
;  CURRENT FILING DATE:  2002-03-02
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 3281
;   TYPE: DNA
;   ORGANISM: human
US-10-090-238-1

  Query Match             91.6%;  Score 432.4;  DB 12;  Length 3281;
  Best Local Similarity   98.9%;  Pred. No. 2e-124;
  Matches  467;  Conservative    0;  Mismatches    1;  Indels    4;  Gaps    3;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1261 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 1320

Qy         61 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db       1321 GGCATCATCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGA-CCTGCTGGG 1379

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1380 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAG 1439

Qy        181 CAGAGGTCAGCACTGGGG-AAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCT 239
              |||||||||||||||||| |||||||||||| | ||||||||||||||||||||||||
Db       1440 CAGAGGTCAGCACTGGGGAAAGACAGGTTGTC--CACAGGGGATGGGGGTCCATCCACCT 1497

Qy        240 TGCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAG 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1498 TGCCGAAAAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAG 1557

Qy        300 AGGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCC 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1558 AGGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCATGACCC 1617

Qy        360 ACGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGAC 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1618 ACGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGAC 1677

Qy        420 AGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1678 AGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAAGC 1729


RESULT 14
US-10-212-667-2
; Sequence 2, Application US/10212667
; Publication No. US20030082722A1
; GENERAL INFORMATION:
;  APPLICANT: FANG, BIANGLIANG
;  TITLE OF INVENTION: METHOD FOR AMPLIFYING EXPRESSION FROM A CELL SPECIFIC
;  TITLE OF INVENTION:  PROMOTER
;  FILE REFERENCE: UTSC:708US
;  CURRENT APPLICATION NUMBER: US/10/212,667
;  CURRENT FILING DATE:  2002-08-05
;  PRIOR APPLICATION NUMBER: 60/310,905
;  PRIOR FILING DATE: 2001-08-08
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 425
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
US-10-212-667-2

  Query Match             74.2%;  Score 350.4;  DB 14;  Length 425;
  Best Local Similarity   98.7%;  Pred. No. 3.7e-99;
  Matches  385;  Conservative    0;  Mismatches    1;  Indels    4;  Gaps    3;

Qy         83 GCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGGTTTCTCTGTCACAAAGGAAAAT 142
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db          1 GCCCTGGAGAGCATGGGGAGACCCGGGA-CCTGCTGGGTTTCTCTGTCACAAAGGAAAAT 59

Qy        143 AATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAGCAGAGGTCAGCACTGGGG-AAG 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db         60 AATCCCCCTGGTGTGACAGACCCAAGGACAGAACACAGCAGAGGTCAGCACTGGGGAAAG 119

Qy        202 ACAGGTTGTCCTCCCAGGGGATGGGGGICCATCCACCTTGCCGAAAAGATTTGTCTGAGG 261
              |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Db        120 ACAGGTTGTC--CACAGGGGATGGGGGTCCATCCACCTTGCCGAAAAGATTTGTCTGAGG 177

Qy        262 AACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGAAATGAGAGGGGAGGG 321
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        178 AACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGAAATGAGAGGGGAGGG 237

Qy        322 GACAGAGGACACCTGAATAAAGACCACACCCATGACCCACGTGATGCTGAGAAGTACTCC 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        238 GACAGAGGACACCTGAATAAAGACCACACCCATGACCCACGTGATGCTGAGAAGTACTCC 297

Qy        382 TGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCA 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        298 TGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCA 357

Qy        442 GCCTTGACAAAACGTTCCTGGAACTCAAGC 471
              ||||||||||||||||||||||||||||||
Db        358 GCCTTGACAAAACGTTCCTGGAACTCAAGC 387


RESULT 15
US-10-027-632-97812
; Sequence 97812, Application US/10027632
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97812
;   LENGTH: 1855
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-97812

  Query Match           25.6%;  Score 120.8;  DB 13;  Length 1855;
  Best Local Similarity  66.1%;  Pred. No. 2.2e-27;
  Matches  292;  Conservative    0;  Mismatches  122;  Indels   28;  Gaps    7;

Qy         18 CCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTGGGCATCATCCCACCTTC 77
              || ||||||| ||| ||||||||||||||||||||||||| |||| |||||| |||  ||
Db         95 CCATTTTCTATCCCACAGAGCCACCTCTGTCACCTTCCTGCTGGGTATCATCTCACACTC 154

Qy         78 CCAGAGCCCTGGAGAGCATGGGGAGACC-CGGGACCCTGCTGGGTTTCTCTGTCACAAAG 136
              || |||   ||| ||||||| |||||||  ||| ||| ||||||| ||| |||||||||
Db        155 CCTGAGTATTGGGGAGCATGAGGAGACCTGGGGGCCCAGCTGGGTCTCTGTGTCACAAAA 214

Qy        137 GAAAATAATCCCCCTGGTGTGACAGACCCAAGGACA--------------GAACACAGCA 182
              | ||| | | ||||  || ||  |||||| ||   |              ||      ||
Db        215 GGAAACAGTTCCCCAAGTTTGGGAGACCCCAGAGTACCTCTGTTTGTGGTGACATTCCCA 274

Qy        183 GAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGG--TCCATCCACCTT 240
               |||         | |  |||| ||   |||| | |||||  ||||  |||| | ||| |
Db        275 AAGGGTCAGTGCAGAGGTGACAAGTCACCCTCTCTGGGGACAGGGGACTCCACCAACCCT 334

Qy        241 GCCGAAAAGATTTGTCTGAGGA------ACTGAAAATAGAAGGGAAAAAAG-AGGAGGGA 293
              ||     | | |    | ||||      | || | | ||||| ||||   | ||||||||
Db        335 GCTTCTCAAAGTGTGGTTAGGAAACTGTAATGTACACAGAAGAGAAAGGGGAAGGAGGGA 394

Qy        294 CAAAAGAGGCAGAAATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGACCACACCCA 353
              ||||| |||||||||||||||||||||||     ||    |     ||   || | ||

Db        395 CAAAAAAGGCAGAAATGAGAGGGGAGGGGCAGAGGGGTGACCTGGGAAGAGCCCCGCCTC 454

Qy        354 TGACCCACGTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGG 413
              || |||   ||   ||| ||||| || ||||||  ||| ||| |||||  |||| |||||
Db        455 TGCCCC---TGGCCCTGGGAAGTGCTTCTGCCC-GGGAGGAGGCTCAGCACAGAAGGAGG 510

Qy        414 AAGGACAGCAGACCAGACAGTC 435
              |||| |||||| || |||||
Db        511 AAGGTCAGCAGCCCCGACAGCC 532


Search completed: September 26, 2003, 14:38:44
Job time : 209 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: September 26, 2003, 12:16:11 ; Search time 2041 Seconds
(without alignments)
5620.633 Million cell updates/sec

Title: US-10-045-116-1
Perfect score: 472
Sequence: 1 AGCCACCACCCAGTGAGCCT.....ACGTTCCTGGAACTCAAGCA 472

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	87.8	18.6	328	9	AW449279	AW449279 UI-H-BI3-
2	75.2	15.9	522	28	AZ553789	AZ553789 RPCI-23-1
3	71	15.0	588	28	AZ377232	AZ377232 1M0131P23
4	69	14.6	595	9	AL602851	AL602851 DKFZp686P
5	68.8	14.6	787	9	AU140869	AU140869 AU140869
6	67	14.2	220	12	BM829043	BM829043 K-EST0101
7	67	14.2	275	12	BM836430	BM836430 K-EST0112
8	67	14.2	279	12	BM752482	BM752482 K-EST0029
9	67	14.2	355	12	BM836104	BM836104 K-EST0111
10	67	14.2	418	12	BM751932	BM751932 K-EST0028
11	67	14.2	483	12	BM828560	BM828560 K-EST0101
12	67	14.2	569	12	BM836866	BM836866 K-EST0112
13	64.6	13.7	534	12	BM752131	BM752131 K-EST0028
14	63	13.3	879	28	AQ746357	AQ746357 HS_2278_A
c 15	62	13.1	195	9	AW839587	AW839587 RC6-LT007
16	61	12.9	651	9	AL602396	AL602396 DKFZp686F
17	59	12.5	390	9	AA132598	AA132598 zo17c08.r
18	57.4	12.2	574	2	HSM069849	Bx479959 Homo sapi
19	56	11.9	887	12	BI759915	BI759915 603044336
20	55	11.7	380	9	AA132182	AA132182 zo28b09.r
c 21	53	11.2	997	29	CNS005TE	AL060767 Drosophil
c 22	52	11.0	506	13	BQ561499	BQ561499 H4070F10-
23	52	11.0	532	12	BM836657	BM836657 K-EST0112
c 24	52	11.0	639	12	BM236450	BM236450 K0435E12-
25	52	11.0	736	9	AL702861	AL702861 DKFZp686B
26	51.6	10.9	300	9	AU099952	AU099952 AU099952
27	50	10.6	50	9	AU105569	AU105569 AU105569
28	50	10.6	835	14	CB995910	CB995910 AGENCOURT
29	50	10.6	930	14	CB989395	CB989395 AGENCOURT
c 30	50	10.6	1101	29	CNS0172T	AL107519 Drosophil
c 31	50	10.6	1155	13	BX444199	BX444199 BX444199
32	49.8	10.6	882	13	BX445657	BX445657 BX445657
c 33	49.8	10.6	985	13	BX359916	BX359916 BX359916
34	49.2	10.4	691	2	HSM069861	Bx479971 Homo sapi
35	49	10.4	523	2	HSM069649	Bx479759 Homo sapi
36	49	10.4	1020	10	BG753754	BG753754 602732959
37	48.8	10.3	595	2	HSM068872	Bx479086 Homo sapi
38	48.4	10.3	768	14	CB994751	CB994751 AGENCOURT
39	48.4	10.3	772	14	CB961513	CB961513 AGENCOURT
40	48.4	10.3	784	10	BG623391	BG623391 602648437
41	48.4	10.3	821	14	CB997421	CB997421 AGENCOURT
42	48.4	10.3	843	14	CB997034	CB997034 AGENCOURT
43	48.4	10.3	847	14	CB989458	CB989458 AGENCOURT
44	48.4	10.3	871	14	CB991637	CB991637 AGENCOURT
45	48.4	10.3	989	14	CB995365	CB995365 AGENCOURT

ALIGNMENTS

RESULT 1
AW449279
LOCUS AW449279 328 bp mRNA linear EST 17-FEB-2000
DEFINITION UI-H-BI3-akh-e-09-0-UI.s1 NCI_CGAP_Sub5 Homo sapiens cDNA clone IMAGE:2734193 3', mRNA sequence.
ACCESSION AW449279
VERSION AW449279.1 GI:6990055
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 328)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward

```
           POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .328
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2734193"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub5"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; NCI_CGAP_Sub5
                     is a subtracted library derived from NCI_CGAP_Sub4. The
                     NCI_CGAP_Sub5 library had 3 million recombinants. A
                     single-stranded DNA preparation of NCI_CGAP_Sub4 was used
                     as a tracer in a subtractive  hybridization with a driver
                     comprising: the IMAGE pool (NCI_CGAP_Kid3 pool 1 LLAM
                     3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs
                     1322376-1323911, 1456008-1456775,1500552-1502855);
                     NCI_CGAP_Kid5 pool 1 LLAM 3338-3342,3722-3725, 3776-3778
                     (IMAGE CloneIDs 1323912-1325831, 1471368-1472903,
                     1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582,
                     3851-3854 (IMAGE  CloneIDs 1414920-1417991,1520904-1522439
                     ); NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720,
                     3733-3735 (IMAGE CloneIDs 1257096-1258631,1469064-1470983,
                     1475592-1476743); NCI_CGAP_Pr22  pool 1 LLAM 2457-2459,
                     2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759
                     ,1101192-1101959,  1217928-1220615); NCI_CGAP_Co10 pool 1
                     LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs  1057416-1061255
                     ,1144584-1145351). (10% of the driver population), plus a
                     pool of 3,840  arrayed clones from  NCI_CGAP_Sub1 (IMAGE
                     CloneIDs 2708616-2710535) and  NCI_CGAP_Sub2 (IMAGE
                     CloneIDs 2710536-2712455)   (10% of the driver population
                     ), plus a pool of 11,136 clones from NCI_CGAP_Sub3 (IMAGE
                     CloneIDs 2712456-2723591) (10% of the driver population),
                     plus a pool of 5,472 clones from NCI_CGAP_Sub4 (IMAGE
                     CloneIDs 2723592-2728969) (70% of the driver population).
                     Subtraction was performed as previously described [Bonaldo
                     , Lennon & Soares (1996): Normalization and Subtraction:
                     Two Approaches To Facilitate Gene Discovery. Genome
                     Research 6, 791-806.
                     TAG_LIB=NCI_CGAP_Co10
                     TAG_TISSUE=colon
                     TAG_SEQ=AAACG"
BASE COUNT       63 a     96 c     80 g     89 t
ORIGIN

  Query Match             18.6%;  Score 87.8;  DB 9;  Length 328;
  Best Local Similarity   65.5%;  Pred. No. 8e-09;
  Matches  154;  Conservative    0;  Mismatches   62;  Indels   19;  Gaps    1;

Qy          1 AGCCACCACCCAGTGAGCCTTTTTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTG 60
              || || | |||||||||||  |||| | ||||||||||||||||||||||||||||| ||
Db         35 AGTCATCCCCCAGTGAGCCCCTTTCCATCCCCCAGAGCCACCTCTGTCACCTTCCTGCTG 94

Qy         61 GGCATCATCCCCACCTTCCCAGAGCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGG 120
              ||  || ||| |||||||||||  | ||| ||||||||||||||||| || | |   ||| |
Db         95 GGTGTCTTCCTACCTTCCCAGCACACTGAAGAGCATGGGGAGACCTGGCAGCTCACTGTG 154

Qy        121 TTTCTCTGTCACAAAGGAAAATAATCCCCCTGG----------------TGTGACAG 161
              || |   | | ||  | | | |       |                  ||||||
Db        155 TTGCAAAGAAAATAATTACCACATTTGGATATGCCAGGGTTCTCTGTGTGCTGCTGACAG 214

Qy        162 ACCCAAGGACAGAACACAGCAGAGGTCAGCACTGGGGAAGACAGGTTGTCCTCCC 216
              |||||  | ||| |||||  |||||||||  |||| || |   | |  |||||||
Db        215 ACCCATAGCCAGGACACAATAGAGGTCAGTGCTGGAGAGGGGGGATCCTCCTCCC 269


RESULT 2
AZ553789
LOCUS       AZ553789                 522 bp    DNA     linear   GSS 20-NOV-2000
DEFINITION  RPCI-23-177H11.TJ RPCI-23 Mus musculus genomic clone RPCI-23-177H11
            , genomic survey sequence.
ACCESSION   AZ553789
            .
            .
VERSION     AZ553789.1  GI:11233609
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
            ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished
COMMENT     Other_GSSs: RPCI-23-177H11.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 177  row: H  column: 11
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .522
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-177H11"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and  partially digested
                     with a combination of EcoRI and EcoRI Methylase.   Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI  sites.  The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies).  "
BASE COUNT      135 a    119 c    136 g    132 t
ORIGIN

  Query Match             15.9%;  Score 75.2;  DB 28;  Length 522;
  Best Local Similarity   55.9%;  Pred. No. 4e-06;
  Matches  237;  Conservative    0;  Mismatches  163;  Indels   24;  Gaps    4;

Qy         23 TTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTGGGCATCATCCCCACCTTCCCAGA 82
              ||| |  |  ||  | ||||| |||||||||||||| |||| |  ||||||||||||||||
Db         92 TTCCAAACTTCACTGGCACCTGTGTCACCTTCCTGCTGGGAA--ATCCCACCTTCCCAGA 149

Qy         83 GCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGGTTTCTCTGTCACAAAGGAAAAT 142
              |  ||| |||    |||  |     |  | |  |||||| | | |||| ||  |  | ||
Db        150 GTACTGAAGACACAGGGCTGTTTGTGATCTCATCTGGGTCTATGTGTCTCAGGGACACAT 209

Qy        143 AATCCCCCTGGTG----------------TGACAGACCCAAGGACAGAACACAGCAGAGG 186
              |       | | |                |||||||| || |  ||| |  |||||||
Db        210 AGGAAAGATAGAGGTTCCCTCTTTGATGCTGACAGACTCATGTCCAGGAGTCAGCAGATA 269

Qy        187 TCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTTGCCGAA 246
              ||||| |  ||    |  |  ||      ||      |   |    || ||
Db        270 TCAGCGCCTGGATGAATCATTGTTCTCTGAGGGCATGTCGATACTTATCATCTGTGTTGC 329

Qy        247 AAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGA 306
                | | |   |||  ||||   ||| || ||||   |||| | |   | |||||
Db        330 TCAAGGTTTTGTTGGAGAAGAAACATAAAGAGAGAAAATGTGAGGAAGATTTGGGGCAGC 389
```

```
Qy        307 AATGAGAGGGGAGGGGACAGAGGA-----CACCTGAATAAAGACCACACCCATGACCCAC 361
              | |||||| ||| | ||| ||| |          | |   |   |  ||| ||   ||||
Db        390 ACTGAGAGTGGAAGTGACTGAGCAGTGTCGTGGTCACAGACCCCACCACACACAGCCCAT 449

Qy        362 GTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAG 421
                ||| ||| ||||| |||||| ||| || | |||||| |||| ||||||||||||||
Db        450 AGGATTCTGGGAAGTGCTCCTG-CCTGAGAGAAAACTCAGCTCAGAAGGAGGAAGGACAG 508

Qy        422 CAGA 425
              ||||
Db        509 CAGA 512
```

RESULT 3
AZ377232
LOCUS AZ377232 588 bp DNA linear GSS 02-OCT-2000
DEFINITION 1M0131P23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0131P23 F, genomic survey sequence.
ACCESSION AZ377232
VERSION AZ377232.1 GI:10490932
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 588)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0131 row: P column: 23
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 588.

FEATURES Location/Qualifiers
source 1. .588
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0131P23"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 157 a 129 c 162 g 139 t 1 others
ORIGIN

Query Match 15.0%; Score 71; DB 28; Length 588;
Best Local Similarity 55.2%; Pred. No. 3.2e-05;
Matches 234; Conservative 0; Mismatches 166; Indels 24; Gaps 4;

```
Qy         23 TTCTAGCCCCCAGAGCCACCTCTGTCACCTTCCTGTTGGGCATCATCCCACCTTCCCAGA 82
              ||| | |   || || || || |||| | |||||| | || |  | | ||||||||||||
Db         98 TTCCAACTGTCACAGACATCTGTGTCCCTTTCCTGATAGGAA--AGCTCACCTTCCCAGA 155

Qy         83 GCCCTGGAGAGCATGGGGAGACCCGGGACCCTGCTGGGTTTCTCTGTCACAAAGGAAAAT 142
              || |||  ||  | | | ||||  ||  | | ||||||| | | |||||||  |  |  |
Db        156 GCTCTGAGGACAAAGTGCAGACTGGGACCTCAGCTGGGTCTGTGTGTCACAGGGACACGT 215

Qy        143 AATCCCCCTGGTG----------------TGACAGACCCAAGGACAGAACACAGCAGAGG 186
              |       ||| |                |||||||| ||       | |  ||||| | |
Db        216 AGGAATGATGGAGGCTCCCTCTTTGATGCTGACAGACTCATATCTGGGAGTCAGCATATG 275

Qy        187 TCAGCACTGGGGAAGACAGGTTGTCCTCCCAGGGGATGGGGGTCCATCCACCTTGCCGAA 246
              ||||| || ||       | |  ||       ||      |   |    || | |
Db        276 TCAGCCCTTGGATGAGTTGTTGTTCTGTGAGGGCATGCCGATGCTTATCAGCCTTGTTGC 335

Qy        247 AAGATTTGTCTGAGGAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGA 306
                |  | |    |||   |||     ||| || ||||   |||  | |   || ||||
Db        336 TCAAGGTCTTGTTGGAGGAGAATCATAAAGAGAGAAAATGTGAGAAAGATGTGATGCAGC 395

Qy        307 AATGAGAGGGGAGGGGACAGAGGACACCTGAATAAAGA-----CCACACCCATGACCCAC 361
              | |||||| ||| | |||  || |  ||     | | |     |  ||| ||   ||||
Db        396 ACTGAGAGTGGAAGTGACTTAGCAGTTCTATGGACACAGACCCCACCACACACAGCCCAT 455

Qy        362 GTGATGCTGAGAAGTACTCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAG 421
              ||| ||| ||||| |||||| |||  ||  | ||||||| || | ||||||||||||||
Db        456 AGGATTCTGGGAAGTGCTCCTG-CCTGAGAGAACACTCAGGTCANAAGGAGGAAGGACAG 514

Qy        422 CAGA 425
              |||
Db        515 TAGA 518
```

RESULT 4
AL602851
LOCUS AL602851 595 bp mRNA linear EST 14-AUG-2001
DEFINITION DKFZp686P1216_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone DKFZp686P1216 5', mRNA sequence.
ACCESSION AL602851
VERSION AL602851.1 GI:15166357
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 595)
AUTHORS Wambutt,R., Heubner,D., Mewes,W., Weil,B. and Wiemann,S.
TITLE EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished
COMMENT Contact: Wambutt R
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de; sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
No s1 sequence available.
This clone (DKFZp686P1216) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES Location/Qualifiers
source 1. .595
/organism="Homo sapiens"

```
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686P1216"
                     /tissue_type="human skeletal muscle"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
BASE COUNT      169 a    178 c    138 g    110 t
ORIGIN

  Query Match             14.6%;  Score 69;  DB 9;  Length 595;
  Best Local Similarity   100.0%;  Pred. No. 8.6e-05;
  Matches   69;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        403 GCAGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGG 462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGG 60

Qy        463 AACTCAAGC 471
              |||||||||
Db         61 AACTCAAGC 69


RESULT 5
AU140869
LOCUS       AU140869                 787 bp    mRNA    linear   EST 05-AUG-2002
DEFINITION  AU140869 PLACE4 Homo sapiens cDNA clone PLACE4000399 5', mRNA
            sequence.
ACCESSION   AU140869
VERSION     AU140869.1  GI:11002390
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 787)
  AUTHORS   Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
            Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.
  TITLE     HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
            Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
            Masuho,Y., Isogai,T.)
  JOURNAL   Unpublished
COMMENT     Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3975
            Fax: 81-438-52-3986
            Email: genomics@hri.co.jp
            HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
            Research Institute; cDNA library construction: Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.
FEATURES             Location/Qualifiers
     source          1. .787
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="PLACE4000399"
                     /tissue_type="placenta"
                     /clone_lib="PLACE4"
                     /note="Vector: pME18SFL3"
BASE COUNT      232 a    205 c    203 g    144 t      3 others
ORIGIN

  Query Match             14.6%;  Score 68.8;  DB 9;  Length 787;
  Best Local Similarity   61.9%;  Pred. No. 9.4e-05;
  Matches  130;  Conservative    0;  Mismatches   72;  Indels    8;  Gaps    1;

Qy        262 AACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGAAATGAGAGGGGAGGG 321
              || || | | |||||||| ||| || || |   ||| |||||| || ||||||||||||||

Db          2 AAATGTAGACAGAAGGGAGAAAGGAAGAAGAGAAAATGAGGCAAAACTGAGAGGGGAGGG 61

Qy        322 GACAGAGGACACCTGAATAAAGACCACACCCATGACCCACGTGATGCTGAGAAGTACTCC 381
              |||||||             ||| |    ||        | |||  |||  |||| ||||

Db         62 GACAGAGAGGTGACCTGGGCAGAGCTTCACC--------CATGACCCTGGAAAGTGCTCC 113

Qy        382 TGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCA 441
              |||||| ||| ||| |||||    ||  |||||||||||||||| |  |

Db        114 TGCCCTGGGAGGAGGCTCAGCATGGAAAGAGGAAGGACAGCAGAGCCTAAGTCACAGTAG 173

Qy        442 GCCTTGACAAAACGTTCCTGGAACTCAAGC 471
               |||    | | | |||||||| | || ||

Db        174 CCCTGACTACAGCATTCCTGGAGCCCAGGC 203


RESULT 6
BM829043
LOCUS       BM829043                 220 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0101950 S9SNU601 Homo sapiens cDNA clone S9SNU601-52-B11 5',
            mRNA sequence.
ACCESSION   BM829043
VERSION     BM829043.1  GI:19185452
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 220)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 52  row: B  column: 11
            High quality sequence stop: 220.
FEATURES             Location/Qualifiers
     source          1. .220
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S9SNU601-52-B11"
                     /sex="M"
                     /tissue_type="Ascites"
                     /cell_type="Epithelial"
                     /cell_line="SNU-601"
                     /lab_host="Top10F'"
                     /clone_lib="S9SNU601"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
BASE COUNT       58 a     77 c     51 g     34 t
ORIGIN

  Query Match             14.2%;  Score 67;  DB 12;  Length 220;
```

```
  Best Local Similarity   100.0%;  Pred. No. 0.00024;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        405 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 60

Qy        465 CTCAAGC 471
              |||||||
Db         61 CTCAAGC 67


RESULT 7
BM836430
LOCUS       BM836430                275 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0112089 S9SNU601 Homo sapiens cDNA clone S9SNU601-61-B06 5',
            mRNA sequence.
ACCESSION   BM836430
VERSION     BM836430.1  GI:19192839
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 275)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 61  row: B  column: 06
            High quality sequence stop: 275.
FEATURES             Location/Qualifiers
     source          1. .275
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S9SNU601-61-B06"
                     /sex="M"
                     /tissue_type="Ascites"
                     /cell_type="Epithelial"
                     /cell_line="SNU-601"
                     /lab_host="Top10F'"
                     /clone_lib="S9SNU601"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
BASE COUNT       71 a     90 c     67 g     47 t
ORIGIN

  Query Match             14.2%;  Score 67;  DB 12;  Length 275;
  Best Local Similarity   100.0%;  Pred. No. 0.00024;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        405 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 60

Qy        465 CTCAAGC 471
              |||||||
Db         61 CTCAAGC 67


RESULT 8
BM752482
LOCUS       BM752482                279 bp    mRNA    linear   EST 04-MAR-2002
DEFINITION  K-EST0029023 S9SNU601 Homo sapiens cDNA clone S9SNU601-6-B10 5',
            mRNA sequence.
ACCESSION   BM752482
VERSION     BM752482.1  GI:19082100
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 279)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 6  row: B  column: 10
            High quality sequence stop: 279.
FEATURES             Location/Qualifiers
     source          1. .279
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S9SNU601-6-B10"
                     /sex="M"
                     /tissue_type="Ascites"
                     /cell_type="Epithelial"
                     /cell_line="SNU-601"
                     /lab_host="Top10F'"
                     /clone_lib="S9SNU601"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
BASE COUNT       72 a     90 c     68 g     49 t
ORIGIN

  Query Match             14.2%;  Score 67;  DB 12;  Length 279;
  Best Local Similarity   100.0%;  Pred. No. 0.00024;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        405 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 60
```

```
Qy        465 CTCAAGC 471
              |||||||
Db         61 CTCAAGC 67


RESULT 9
BM836104
LOCUS       BM836104                355 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0111597 S9SNU601 Homo sapiens cDNA clone S9SNU601-65-F08 5',
            mRNA sequence.
ACCESSION   BM836104
VERSION     BM836104.1  GI:19192513
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 355)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 65  row: F  column: 08
            High quality sequence stop: 355.
FEATURES             Location/Qualifiers
     source          1. .355
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S9SNU601-65-F08"
                     /sex="M"
                     /tissue_type="Ascites"
                     /cell_type="Epithelial"
                     /cell_line="SNU-601"
                     /lab_host="Top10F'"
                     /clone_lib="S9SNU601"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
BASE COUNT       95 a    105 c     87 g     68 t
ORIGIN

  Query Match             14.2%;  Score 67;  DB 12;  Length 355;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        405 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 60

Qy        465 CTCAAGC 471
              |||||||
```

```
Db         61 CTCAAGC 67


RESULT 10
BM751932
LOCUS       BM751932                418 bp    mRNA    linear   EST 04-MAR-2002
DEFINITION  K-EST0028271 S9SNU601 Homo sapiens cDNA clone S9SNU601-18-E04 5',
            mRNA sequence.
ACCESSION   BM751932
VERSION     BM751932.1  GI:19081550
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 418)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 18  row: E  column: 04
            High quality sequence stop: 418.
FEATURES             Location/Qualifiers
     source          1. .418
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S9SNU601-18-E04"
                     /sex="M"
                     /tissue_type="Ascites"
                     /cell_type="Epithelial"
                     /cell_line="SNU-601"
                     /lab_host="Top10F'"
                     /clone_lib="S9SNU601"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
BASE COUNT      115 a    124 c    100 g     79 t
ORIGIN

  Query Match             14.2%;  Score 67;  DB 12;  Length 418;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        405 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 60

Qy        465 CTCAAGC 471
              |||||||
Db         61 CTCAAGC 67
```

```
RESULT 11
BM828560
LOCUS       BM828560                483 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0101388 S9SNU601 Homo sapiens cDNA clone S9SNU601-53-D01 5',
            mRNA sequence.
ACCESSION   BM828560
VERSION     BM828560.1  GI:19184969
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 483)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 53  row: D  column: 01
            High quality sequence stop: 483.
FEATURES             Location/Qualifiers
     source          1. .483
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S9SNU601-53-D01"
                     /sex="M"
                     /tissue_type="Ascites"
                     /cell_type="Epithelial"
                     /cell_line="SNU-601"
                     /lab_host="Top10F'"
                     /clone_lib="S9SNU601"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
BASE COUNT      139 a    144 c    109 g     91 t
ORIGIN

  Query Match           14.2%;  Score 67;  DB 12;  Length 483;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        405 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 60

Qy        465 CTCAAGC 471
              |||||||
Db         61 CTCAAGC 67


RESULT 12
BM836866
LOCUS       BM836866                569 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0112771 S9SNU601 Homo sapiens cDNA clone S9SNU601-78-D07 5',
            mRNA sequence.
ACCESSION   BM836866
VERSION     BM836866.1  GI:19193275
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 569)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 78  row: D  column: 07
            High quality sequence stop: 569.
FEATURES             Location/Qualifiers
     source          1. .569
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S9SNU601-78-D07"
                     /sex="M"
                     /tissue_type="Ascites"
                     /cell_type="Epithelial"
                     /cell_line="SNU-601"
                     /lab_host="Top10F'"
                     /clone_lib="S9SNU601"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
BASE COUNT      162 a    172 c    128 g    107 t
ORIGIN

  Query Match           14.2%;  Score 67;  DB 12;  Length 569;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   67;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        405 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 464
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAA 60

Qy        465 CTCAAGC 471
              |||||||
Db         61 CTCAAGC 67


RESULT 13
BM752131
LOCUS       BM752131                534 bp    mRNA    linear   EST 04-MAR-2002
DEFINITION  K-EST0028539 S9SNU601 Homo sapiens cDNA clone S9SNU601-4-E04 5',
            mRNA sequence.
ACCESSION   BM752131
```

```
VERSION     BM752131.1  GI:19081749
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 534)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 4  row: E  column: 04
            High quality sequence stop: 534.
FEATURES             Location/Qualifiers
     source          1. .534
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S9SNU601-4-E04"
                     /sex="M"
                     /tissue_type="Ascites"
                     /cell_type="Epithelial"
                     /cell_line="SNU-601"
                     /lab_host="Top10F'"
                     /clone_lib="S9SNU601"
                     /note="Organ: Stomach; Vector: pME18-FL3; Site_1: XhoI;
                     Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including SfiI
                     site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized with Superscript II using SfiI
                     oligo-dT primer. After first strand synthesis, RNA was
                     degraded by NaOH treatment and cDNA was amplified by PCR
                     reaction. The PCR products were digested with SfiI and
                     cloned into DraIII- digested pME18S-FL3 vector. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
BASE COUNT      154 a    153 c    124 g    103 t
ORIGIN

  Query Match             13.7%;  Score 64.6;  DB 12;  Length 534;
  Best Local Similarity   94.4%;  Pred. No. 0.00076;
  Matches   67;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        401 GGGCAGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCT 460
              | |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 GTGGCGAGGGAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCT 71

Qy        461 GGAACTCAAGC 471
              ||||||||  ||
Db         72 GGAACTCATGC 82


RESULT 14
AQ746357
LOCUS       AQ746357                879 bp    DNA     linear   GSS 19-JUL-1999
DEFINITION  HS_2278_A2_A04_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2278 Col=8 Row=A, genomic survey
            sequence.
ACCESSION   AQ746357
VERSION     AQ746357.1  GI:5533515
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 879)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2278  row: A  column: 8
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 879.
FEATURES             Location/Qualifiers
     source          1. .879
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2278 Col=8 Row=A"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      208 a    197 c    282 g    167 t     25 others
ORIGIN

  Query Match             13.3%;  Score 63;  DB 28;  Length 879;
  Best Local Similarity   68.4%;  Pred. No. 0.0017;
  Matches  121;  Conservative    0;  Mismatches   45;  Indels   11;  Gaps    2;

Qy        261 GAACTGAAAATAGAAGGGAAAAAAGAGGAGGGACAAAAGAGGCAGAAATGAGAGGGGAGG 320
              |||    | | ||||   |||    | |  ||   |  |||||||  | ||||||||||
Db        350 GAAACACAGACAGAAAATAAAGGGAAAGGAGGGACAGGGAGGCAGGACTGAGAGGGGAGG 409

Qy        321 GGAC--AGAGGACACCTGAATAAAGACCACACCCATGACCCACGTGATGCTGAGAAGTAC 378
              ||||  |||||   ||||   | |||| | |||||| |        |||||||| |
Db        410 GGACAGAGAGGTGTCCTGGGCACAGACCCCGCCCATGAGC---------CTGAGAAGTGC 460

Qy        379 TCCTGCCCTAGGAAGAGACTCAGGGCAGAGGGAGGAAGGACAGCAGACCAGACAGTC 435
              ||||||||  |||||||  |||||  || | |||||||||||||||| | | ||||| |
Db        461 TCCTGCCCCGGGAAGAGGCTCAGCACAAAAGGAGGAAGGACAGCACAGCTGACAGCC 517


RESULT 15
AW839587/c
LOCUS       AW839587                195 bp    mRNA    linear   EST 18-MAY-2000
DEFINITION  RC6-LT0072-020200-021-B10 LT0072 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW839587
VERSION     AW839587.1  GI:7933561
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 195)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
```

```
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=RC6-LT0072-020
            200-021-B10&t3=2000-02-02&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 195.
FEATURES             Location/Qualifiers
     source          1. .195
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="LT0072"
                     /note="Organ: leiomios; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       31 a     46 c     67 g     51 t
ORIGIN

  Query Match             13.1%;  Score 62;  DB 9;  Length 195;
  Best Local Similarity   100.0%;  Pred. No. 0.0028;
  Matches   62;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        410 GAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAA 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 GAGGAAGGACAGCAGACCAGACAGTCACAGCAGCCTTGACAAAACGTTCCTGGAACTCAA 136

Qy        470 GC 471
              ||
Db        135 GC 134


Search completed: September 26, 2003, 13:41:12
Job time : 2046 secs
```